UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE FILOSOFÍA Y LETRAS

DEPARTAMENTO DE FILOLOGÍA INGLESA

Programa de Doctorado: Lingüística Aplicada

TESIS DOCTORAL

Assessment for Learning in the CLIL classroom:

A corpus based study of teacher motivational L2 strategies and student motivation and metacognitive abilities

Rachel Lofft Basse

Directora: Dr. Rachel Whittaker Dovey

Madrid, 2016

Abstract

This dissertation investigates the relationship between assessment for learning (AfL) and motivation in Content and Language Integrated Learning (CLIL) classrooms. AfL is "the process of seeking and interpreting evidence for use by learners and their teachers to decide where learners are in their learning, where they need to go and how best to get there" (Broadfoot et al., 2002: 2-3). Research has shown the positive effects of AfL on student motivation (Butler, 1988; Ross, Siegenthaler and Tronson, 2006; Schunk, 1996). However, investigation is needed to determine its effects on English language learners (ELLs) (Alvarez et al., 2014; Heritage et al 2013) and lower achieving learners, which have been cited as groups that can benefit the most from AfL (Boston, 2002). Additionally, assessment has been cited as an under researched area in CLIL (Llinares, Morton and Whittaker, 2012), which this study seeks to address.

Participants come from four bilingual schools located in Madrid, Spain and the surrounding areas. This is a corpus-based, Mixed-Methods study investigating AfL using three perspectives. In the first perspective, 14 classroom recordings (six full didactic units, totaling 71,504 words) taught by four teachers (two AfL and two non-AfL) are analyzed using the Motivated Orientation of Language Teaching (MOLT) scheme (Guilloteaux and Dörnyei, 2008). This perspective seeks to identify L2 motivational strategies found in the lessons and examine the relationship of these strategies to AfL techniques. The frequency and distribution are taken into account, as well as a detailed, manual, real-time analysis of the six units, achieving an in-depth picture of AfL techniques in the classroom context. The second perspective is student-centered and seeks to measure the motivation of a sample group of students (N=40) as well as their feelings in certain classroom situations through a questionnaire. Finally, the third perspective, also student-centered, focuses on interviews with lower achieving students (N=6), which encourage reflections on

their own learning and classroom experience. The interviews are analyzed using Martin and White's Appraisal framework (2005).

The findings show that the implementation of AfL coincided with an increased frequency and variety of L2 motivational strategies used by the teacher. This led to classes in which activities conducted were supported by a more motivational discourse. However, no significant differences were seen in the motivation of AfL and non-AfL students. Additionally, AfL students expressed more uncertainty while self-assessing their own abilities. Nevertheless, the results indicate a tendency in AfL classes to promote active learning and present concrete aims, which are reflected upon through peer and self-assessment. The effects of these techniques were seen in the responses of lower achieving AfL students, who were more critical and detailed in their self-assessment than their non-AfL counterparts. This is one of the first studies to examine the effects of AfL in CLIL empirically using a corpus. The findings of this study contribute to an understanding of how the use of AfL techniques would make it possible to integrate motivational strategies in a more systematic way in the CLIL educational context.

Key words: Assessment for learning (AfL), Content and Language Integrated Learning (CLIL), motivation, lower achieving students, Appraisal

Acknowledgements

This dissertation would have never been completed without a tremendous amount of support from a number of people. I would like to take this opportunity to express my gratitude.

First and foremost, I am deeply grateful to my supervisor, Dr. Rachel Whittaker, who encouraged me to enroll in this doctoral program and provided endless support, patience and inspiration throughout this process. Little did I know when I walked into her office for the first time that I would be meeting a mentor who would be a steadfast source of encouragement and guidance for the next seven years. Thank you for always keeping me motivated with three simple words at the end of each meeting, "On we go!"

My utmost appreciation goes to Dr. Christiane Dalton-Puffer (Universität Wien) and Dr. Mary Schleppegrell (University of Michigan), my supervisors during my research stays abroad. Their insightful feedback and support allowed me to gain a fresh perspective and make progress on my work.

I am very thankful to the Universidad Autónoma de Madrid for making all of this possible by providing me with the funding to pursue my doctoral studies full time with the *Inicio de Estudios* and *Formación Personal Investigador (FPI)* grants. Additionally, I am grateful for the special mobility grants to continue this research in different university contexts and attend research conferences. The experience that I have gained here has been life changing and I will be forever grateful.

I am especially grateful to all of the amazing faculty members at the Universidad Autónoma de Madrid who have helped me immensely from the beginning. Thanks to Dr. Ana Llinares for allowing me to be a part of the UAM-CLIL research team and for her expertise in the design of the study and data collection, which shaped the method of this dissertation. To Dr. Mick O'Donnell, whose office door was always open to my questions regarding the UAM CorpusTool and the interpretation of the statistics. To Dr. Clara Molina for her willingness to help in the grant application process and for her encouragement all of these years. To Dr. Tom Morton for

helping me navigate the initial stages of Appraisal analysis and for always giving such insightful feedback. A very special thank you to Dr. Luisa Martín Rojo and Dr. Héctor Grad for their guidance at the beginning of this process with the data collection procedure and the adaptation of the questionnaires.

My gratitude also goes to Dr. Teppo Jakonen (University of Helsinki), Dr. Oliver Meyer (Johannes Gutenberg-University/Mainz) and Dr. Liss Kerstin Sylvén (University of Gothenburg) for their time and willingness in providing external reviews of this dissertation.

A special thanks to Dr. Lucía Poggio (Universidad Complutense) for her help with the Spanish translations.

I am particularly grateful to all of the teachers and students from schools involved in this project. Thank you for inviting me into your classrooms and for allowing me to observe and videotape your daily work. Without your dedication and collaboration, this study would not have been possible.

A very special thank you to Irene Pascual, my partner throughout this process. From the collection of the corpus to participation in conferences and projects to sharing daily trials and tribulations, you have been there every step of the way. I am so grateful that we have shared this journey together.

My thanks also go to my fellow Ph.D. and Masters students from the UAM, who offered encouragement and endless coffee at Juanjo's. To my Ph.D. student friends from other disciplines for being sounding boards throughout this process and to my friends for always offering their support.

Finally, I would like to thank my incredible family and extended family, who have encouraged me not only in the pursuit of this doctorate, but in the pursuit of all my dreams. Last but not least, my deepest gratitude goes to my husband, Alfonso, for being by my side in every moment of doubt and inspiration.

Spanish Presentation

LA EVALUACIÓN PARA EL APRENDIZAJE EN AICLE:

EL ESTUDIO DE LAS ESTRATEGIAS MOTIVACIONALES 2L DE PROFESORES Y LA MOTIVACIÓN Y CAPACIDAD METACOGNITIVA DE LOS ALUMNOS A TRAVÉS DE UN CORPUS

This section presents some parts of the thesis in Spanish. First, a brief summary of the dissertation is provided. The presentation continues with an introduction and ends with the conclusions obtained.

Esta sección presenta algunas partes de la tesis en español. En primer lugar, se describe brevemente la tesis mediante un resumen. Luego, se centra en la introducción del estudio. Los resultados y las implicaciones se mencionan al final.

Resumen

En la presente tesis se estudia la relación entre la evaluación para el aprendizaje (*Assessment for Learning* o AfL) y la motivación en clases de Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (AICLE). AfL es "el proceso de búsqueda e interpretación de evidencias usadas por los alumnos y sus profesores con el objetivo de decidir dónde están los alumnos en el proceso de aprendizaje, hacia dónde necesitan ir y cuál es la mejor manera para llegar allí" (Broadfoot et al., 2002: 2-3). Los estudios realizadas han mostrado los efectos positivos de la AfL en la motivación de los alumnos (Butler, 1998; Ross, Siegenthaler y Tronson, 2006; Schunk, 1996). Sin embargo, son necesarias estudios que determinen sus efectos en alumnos que estudian inglés como lengua extranjera (Alvarez et al., 2014; Heritage et al., 2013) y en alumnos de bajo rendimiento; dos de los grupos que más pueden beneficiarse de la AfL (Boston, 2002). Asimismo, se ha identificado la evaluación como una de las áreas menos estudiados dentro del AICLE (Llinares, Morton y Whittaker, 2012), y que se pretende abordar con este estudio.

Los participantes en la presente tesis provienen de cuatro colegios bilingües localizados en Madrid (España) y sus alrededores. Se ha realizado un estudio de métodos mixtos basado en un corpus y que estudia AfL según tres enfoques distintos. Para el primer enfoque, se grabó un corpus de 14 clases (compuesto por seis unidades didácticas, con 71,504 palabras en total), impartidas por cuatro profesores (dos AfL y dos no-AfL), que está analizado según el esquema *Motivated Orientation of Language Teaching* (MOLT) (Guilloteaux y Dörnyei, 2008). Este enfoque busca identificar las estrategias de motivación en relación a la segunda lengua (L2) presentes en cada lección para examinar la relación entre estas estrategias y las técnicas AfL. Se ha estudiado la frecuencia y la distribución de las estrategias, así como un análisis temporal detallado de las seis unidades didácticas, lográndose una imagen del uso de la AfL en las aulas. El segundo enfoque está centrado en el alumno y tiene como objetivo, a través de un cuestionario, medir tanto la motivación del grupo de muestra de alumnos (N=40) como sus sentimientos en varias situaciones en el aula. Finalmente, el tercer enfoque también está orientado hacia el alumno y se basa en entrevistas con alumnos de bajo rendimiento (N=6) que reflexionan acerca de su propio aprendizaje y de su experiencia en el aula. Las entrevistas están analizadas usando el esquema de lenguaje evaluativo (*Appraisal*) de Martin y White (2005).

Los resultados muestran que la implementación de la AfL coincide con un incremento de la frecuencia y una variedad de estrategias motivacionales L2 usadas por los profesores. Además, en estas clases las actividades están apoyadas por un discurso más motivacional. Sin embargo, no se han encontrado diferencias significativas entre la motivación de alumnos de AfL y no AfL. Por otro lado, se observa una mayor tendencia de los alumnos de AfL a la hora de reflexionar acerca de sus habilidades. No obstante, los resultados indican una tendencia en clases de AfL para promover el aprendizaje activo y presentar objetivos concretos, que se reflejan a través de la evaluación en pareja y la autoevaluación. Los resultados de estas técnicas se pudieron observar en las respuestas de los alumnos de bajo rendimiento en clases de AfL, quienes eran más críticos en su autoevaluación que

los alumnos de no-AfL. Este es uno de los primeros estudios que examinan empíricamente los efectos del AfL en el contexto de AICLE usando un corpus. Los resultados de este estudio contribuyen a la comprensión de cómo el uso del AfL puede facilitar la integración de estrategias motivacionales de una forma más sistemática en el contexto de AICLE.

Palabras claves: Evaluación para el Aprendizaje (AfL), Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (AICLE), motivación, alumnos de bajo rendimiento, lenguaje evaluativo (*Appraisal*)

Introducción

"La evaluación es un regalo que damos a nuestros alumnos. Es un espejo que damos para enseñarles su progreso. La evaluación es una promesa que vamos a utilizar, no para castigar ni premiar, sino para guiarles en su camino de aprendizaje."

Jan Chappuis, Assessment Training Institute

Motivación del estudio

La investigación se centra en el uso de la Evaluación para el Aprendizaje (AfL) y su relación con la motivación en contextos de Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (AICLE). La definición de la AfL según el grupo de Reforma de la Evaluación es:

> (la AfL es) El proceso de búsqueda e interpretación de evidencias usadas por los alumnos y sus profesores con el objetivo de decidir dónde están los alumnos en el proceso de aprendizaje, hacia dónde necesitan ir y cuál es la mejor manera para llegar allí (Broadfoot et al., 2002: 2-3).

Una gran parte de los estudios se ha enfocado en los beneficios de la evaluación formativa, o la evaluación para el aprendizaje. Sin embargo, existen pocas investigaciones que hayan examinado los efectos que se producen en alumnos que están aprendiendo inglés como una segunda lengua (Alvarez et al., 2014). La necesidad de estudios en este área ha sido destacada por Heritage et al. (2013) para poder determinar el potencial de la evaluación y cumplir con las necesidades de contenido y lenguaje.

> Con la expansión de la evaluación formativa, es necesario examinar el impacto potencial del uso de la evaluación formativa tanto en el rendimiento académico como en el aprendizaje de idiomas de los alumnos, así como explorar cómo el uso de la evaluación formativa se puede adaptar a las necesidades especificas de estos alumnos (2013).

Este estudio tiene por objeto colmar esa laguna, investigando la AfL en el contexto de clases AICLE en el cual los alumnos aprenden el lenguaje y el contenido de

manera simultánea. AICLE se refiere a cualquier "actividad en que una lengua extranjera se utiliza como una herramienta en el aprendizaje de una materia no-lingüística y en la que tanto el lenguaje como el contenido tienen un papel conjunto" (Marsh, 2002:58). Es decir, situaciones en las que el contenido se enseña a través de una lengua extranjera, dando la misma importancia a ambos. AICLE es un método de enseñanza introducido en los años 90 que se ha integrado en contextos bilingües de toda Europa (será elaborado en la sección 1.3). A pesar de la implementación de AICLE durante los últimos 25 años, el papel de la evaluación ha sido identificado como un área de AICLE poco estudiado (Llinares, Morton and Whittaker, 2012; Coyle, 2010; Maggi, 2012; Barbero, 2012).

La investigación sobre el papel de la evaluación en AICLE se ha centrado en el lenguaje (Höning, 2009), analizando el nivel lingüístico de los alumnos (Wewer, 2014). El proyecto *Assessment and Evaluation in CLIL* (AECLIL) fue desarrollado para estudiar la evaluación en AICLE (Barbaro, 2012). Con financiación de la Comisión Europea (EAECA), el objetivo de este proyecto es desarrollar recursos didácticos para facilitar la evaluación en clases de AICLE. Una de las mayores preocupaciones en AICLE ha sido la dificultad de alcanzar los objetivos lingüísticos, respecto a la lengua extranjera, debido a la carga lectiva extra resultante de enseñar los contenidos en una lengua extranjera (Clegg, 2012). La evaluación formativa ha sido destacada como un medio para facilitar la integración en ambas áreas (Kiely, 2009; Coyle, 2010), siendo integrada en unos contextos bilingües en España, el contexto en que esta investigación está localizada.

Debido a la presión para alcázar metas lingüísticas y de contenido, la necesidad de motivar a los alumnos de AICLE es alta, especialmente en el caso de alumnos de bajo rendimiento. Tradicionalmente, AICLE ha sido enfocado en alumnos de alto rendimiento (Denman, Tanner and De Graff, 2013), observándose una necesidad de estudiar a alumnos de bajo rendimiento puesto que es un área que requiere más atención (Dobson, Pérez and Johnstone, 2010). La investigación de la motivación en aulas de AICLE se ha centrado en institutos, con pocos estudios dirigidos a colegios

de primaria (Lasagabaster, 2013). Este estudio intenta determinar el potencial de la AfL para promocionar un discurso más motivacional en alumnos de primaria de AICLE, así como para atender a las necesidades de alumnos de bajo rendimiento.

Propuesta y alcance del estudio

Este estudio examina la relación entre la evaluación para el aprendizaje (AfL) y la motivación en aulas de AICLE desde tres perspectivas. La primera se centra en analizar un corpus de unidades didácticas comparando estrategias de motivar a los alumnos, para aprender una segunda lengua, utilizadas por profesores de AfL y no-AfL. La segunda se enfoca en el punto de vista de los alumnos, analizando la autoevaluación de la motivación de alumnos de AICLE en ámbitos AfL y no-AfL. Esta perspectiva también mide los sentimientos de estos alumnos en distintas situaciones en el aula. La tercera perspectiva tiene por objetivo determinar la manera en que los alumnos de bajo rendimiento se autoevalúan y evalúan su ámbito de aprendizaje a través de entrevistas. Hay varios estudios que muestran los efectos positivos de la AfL en la motivación de los alumnos (Ross, Tronson and Siegenenthalter, 2006; Butler, 1988; Schunk, 1996). La mayoría de estos estudios están basados en cuestionarios, entrevistas y los grupos focales. Este estudio usa métodos mixtos para proporcionar una visión más completa de los datos (Johnson, Onwuegbuzie & Turner, 2007), utilizando una combinación de análisis cuantitativo y cualitativo y varios conjuntos de datos.

La mayoría de estudios de la AfL se situan en Reino Unido y los Estados Unidos, dos países en los que el uso de la evaluación formativa está muy integrado en la formación del profesorado. La presente investigación se realiza en Madrid, España, un ámbito importante para el bilingüismo fomentado en contextos escolares. El gobierno de la Comunidad de Madrid ha difundido el bilingüismo en las últimas décadas con el objeto de facilitar el aprendizaje de idiomas desde una edad temprana. El estudio se centra en colegios que pertenecen a programas bilingües en los cuales un 40% del currículo se imparte en inglés. Las siguientes secciones

describen AICLE y su implantación en España para ofrecer un contexto para el estudio.

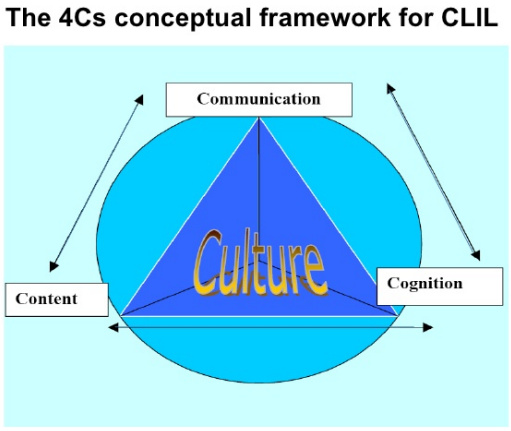
Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (AICLE)

El objetivo de crear una comunidad Europea multilingüe ha sido un proceso continuado durante las últimas dos décadas. La preservación de lenguas Europeas empezó en los años 90, con la publicación del *White Paper on Education and Training* por la Comisión de las Comunidades Europeas, que supuso el compromiso de iniciar desde la escuela infantil la educación bilingüe (Comisión de las Comunidades Europeas, 1995). Debido a que el inglés se ha convertido en una lingua franca dentro de la Comunidad Europea, su emergencia como el idioma principal en estos contextos bilingües comenzó a surgir entre los años 1995 a 2005 (Marsh 2006).

Desde los años 1990, la influencia de AICLE ha sido extensa. Pero, según un informe de Eurydice de 2006, el modelo AICLE está incompleto y requiere más investigación para desarrollar su marco teórico (Dalton-Puffer, 2008). Desde entonces, la investigación académica del AICLE se ha expandido rápidamente por Europa, con la meta de estudiar y definir el marco teórico además de las ventajas y desventajas de su implementación. A través de la realización de varios estudios sobre AICLE dentro de la comunidad Europea, los investigadores están descubriendo los efectos que tiene en el aprendizaje de idiomas.

Además de preparar a los alumnos para ser miembros de una comunidad más internacional, mejorando su competencia comunicativa, AICLE tiene un efecto positivo en la motivación hacia el aprendizaje de idiomas, mejora el conocimiento del vocabulario y permite el aprendizaje incidental, centrándose en la comunicación (Lasagabaster, 2008). El marco de las “4C’s” (Coyle, 1999) facilita los aspectos más importantes de AICLE: el contenido, la comunicación, la cognición y la cultura.

Figura 1.1 Los 4C's de AICLE



Coyle (1999, 2005)

La mayor ventaja de AICLE, que lo distingue de la forma tradicional de aprender inglés como una lengua extranjera, es el enfoque en la comunicación oral entre los profesores y alumnos, que promueve el desarrollo lingüístico (Lasagabaster, 2008). Este énfasis en la comunicación influye en varias competencias lingüísticas, anotados en la tabla 1.1

Tabla 1.1 Competencias lingüísticas impactadas favorablemente o no afectados por AICLE

Impacto positivo	**Sin impacto o indefinido**
Capacidades receptivas	Sintaxis
Vocabulario	Escritura
Morfología	Lenguaje informal o no técnico
La creatividad, asunción de riesgos, fluidez, cantidad	Pronunciación
Resultados emocionales o afectivos	Pragmática

(Dalton-Puffer, 2008: 5)

AICLE mejora las habilidades auditivas, vocabulario (específico de la asignatura), fluidez y habilidad de expresión en la segunda lengua. Además, las investigaciones

han demostrado una influencia positiva en la motivación de alumnos de AICLE (Seikkula-Leino, 2007, Lorenzo et al., 2009; Lasagabaster, 2011; Lasagabaster, 2013; Doiz, Lasagabaster and Sierra, 2014; Sylvén and Thompson, 2015).

A pesar de la multitud de investigaciones sobre las ventajas de AICLE, las críticas se centran en retos tales como: el aislamiento entre profesores de AICLE (Cross and Gearon, 2013); el proceso de selección en programas bilingües que pueden crear una jerarquía entre alumnos AICLE y no-AICLE (Bruton, 2011); y el reto de buscar un equilibrio entre el contenido y el lenguaje (Rowe and Coonan 2008*).* Las limitadas guías existentes sobre CLIL han sido consideradas como un problema, considerando el reto de enfrentarse a necesidades de contenido y lenguaje en el aula (Meyer, 2010). Tal y como se ha mencionado anteriormente, este hecho es verdaderamente relevante en la consideración de la evaluación en AICLE (Coyle, 2010). Como consecuencia, el proyecto AECLIL (Barbaro, 2012) empezó a desarrollar recursos de evaluación para medir la competencia lingüística de alumnos (Barbero, 2012); esquemas para analizar las redacciones y discursos (Barkovsa, 2012); y, finalmente, esquemas para facilitar la autoevaluación y la evaluación de los compañeros (Maggi, 2012). Aunque la elaboración de dichos recursos es necesaria para definir el papel de la evaluación en AICLE, la investigación es imprescindible para poder observar la implementación de la evaluación en aulas de AICLE y determinar los efectos en los alumnos.

La educación bilingüe en España

En las últimas décadas, la implantación de la educación bilingüe ha ido aumentando progresivamente en las Comunidades Autónomas Españolas. Según un estudio de la Comisión Europea, en el año 2006 solo el 17% de ciudadanos españoles reconocieron su habilidad para comunicarse en una lengua extranjera, comparado con el 56% de ciudadanos de la Comunidad Europea. Además, el 56% de españoles admitieron ser monolingües. Por eso, España se quedó en la categoría más baja de competencia lingüística de la Unión Europea junto con países como Italia, Reino Unido e Irlanda (Comisión Europea, 2006). Desde entonces, han proliferado una

multitud de programas AICLE con la creencia de que "representan la mejor manera para mejorar el bajo dominio de lenguas extranjeras visto en alumnos españoles" (Ruiz de Zarobe and Lasagabaster, 2010).

Para intentar mejorar la competencia en una lengua extranjera, España empezó a lanzar iniciativas en los años 90. Estas iniciativas se centraron en el aprendizaje del inglés, que ha sido identificado como uno de los idiomas más comúnmente hablados en Europa (Eurydice, 2006). En el año 1990, se aprobó la *Ley Orgánica de Ordenación General del Sistema Educativo* (una reforma educativa) estipulando la introducción de una lengua extranjera a la edad de ocho años (Gobierno de España, Ministerio de la Presidencia, 1990). Desde entonces, la edad ha bajado hasta los tres años, cuando los niños españoles empiezan su primer año de educación infantil. Este cambio se basa en en investigaciones que afirman que empezar con una edad más temprana puede facilitar el aprendizaje de un idioma (Lasagabaster, 2008; Halbach, 2009).

Como resultado de estos cambios en la política educativa, con un énfasis en el multilingüismo, el British Council junto con el Ministerio de Educación y Ciencias lanzó el Proyecto Bilingüe en el año 1996. Este proyecto implantó el AICLE en centros elegidos en diferentes zonas del país. El proyecto empezó en colegios públicos con la meta de extenderse a institutos antes del año 2004. La iniciativa ha tenido mucho éxito, integrando el aprendizaje de idiomas en el sistema educativo y preparando a las nuevas generaciones para trabajar en otros países Europeos en el futuro (Reilly y Medrano, 2009). Desde el principio, el programa se ha difundido por España y está bien establecido en Madrid. En una evaluación externa en 2010, un grupo de investigadores han comprobado que el proyecto esta alcanzado sus objetivos y que los alumnos "demuestran fluidez y confianza en su dominio del inglés, que incluye vocabulario técnico y la producción de respuestas orales relativamente largas" (Dobson, Pérez y Johnstone, 2010). En 2004, la Comunidad de Madrid (CAM) lanzó un proyecto bilingüe parecido en colegios e institutos públicos de Madrid y sus alrededores. El proyecto se ha extendido en los últimos 10 años, participando hoy en día más que 300 colegios y 100 institutos (Comunidad de

Madrid). Por tanto, el sistema educativo bilingüe en Madrid representa el esfuerzo de promocionar el aprendizaje de lenguas extranjeras en alumnos jóvenes, con más que el 40% de las asignaturas enseñadas en inglés.

Antecedentes de la investigación

El impulso del presente estudio empezó en el año 2009 con un contacto entre el equipo de investigación UAM-CLIL y el British Council en Madrid. El equipo solicitó acceso a varios colegios bilingües para realizar una investigación sobre la evaluación en el aula. Como respuesta, el British Council informó al equipo de una nueva iniciativa por la que se impartían cursos de formación en la evaluación para el aprendizaje a los profesores. Desde entonces, la investigación se ha centrado en estudiar la AfL y sus efectos en los estudiantes del Proyecto Bilingüe. En aquel momento, pocos profesores habían participado en la formación profesional de las técnicas de AfL. El objetivo fue determinar el efecto de éstas técnicas en el aprendizaje de los alumnos, y si la formación de más profesores sería beneficioso. Tras una revisión de las publicaciones específicas de AfL, se decidió centrar el presente estudio en la relación entre la AfL y la motivación, un área que requería más investigación empírica.

Participantes y contexto del estudio

Los participantes procedieron de cuatro colegios de la Comunidad de Madrid, que pertenecieron a los dos programas mencionados anteriormente. Cuatro profesores participaron en el estudio: dos de ellos trabajaron en el proyecto bilingüe del British Council y habían participado en cursos de formación AfL, y dos trabajaron en el programa bilingüe de la Comunidad de Madrid y no tenían formación en técnicas de AfL. Las materias impartidas eran: ciencias naturales, ciencias sociales, ciudadanía, plástica, teatro e inglés. Puesto que este estudio se centra sólo en materias de contenido no incluye la asignatura de inglés. Los alumnos (10-12 años) cursaban 5º y 6º de primaria y empezaron su educación bilingüe a los tres años. Los cuatro colegios están localizados en zonas suburbanas de la Comunidad de Madrid. El nivel socioeconómico era de medio a medio-alto.

Objetivos y preguntas de investigación

La presente investigación usa una variedad de enfoques para estudiar la relación entre la AfL y la motivación. Los datos consisten en un corpus de grabaciones recopilado en el año académico 2010/2011. El corpus representa las siguientes materias: ciencias naturales, ciudadanía, plástica y teatro. Además, el estudio utiliza datos de cuestionarios y entrevistas con una muestra reducida de alumnos.

La estructura del estudio, objetivos y preguntas de investigación son los siguientes:

Parte 1 **(Capítulo 6)**	**Estrategias motivacionales para aprender una segunda lengua en clases de AfL y no-AfL**
Objetivo principal: *Comparar el discurso en clases de AfL y no-AfL, centrándose en las estrategias motivacionales para aprender una segunda lengua (L2)*	
Pregunta de investigación 1	¿Hay una diferencia en la frecuencia y la distribución de las estrategias motivacionales L2 para aprender una segunda lengua dependiendo del uso de AfL?
Pregunta de investigación 2	¿Cómo varia la duración de estas estrategias motivacionales L2 dependiendo de la materia (ciencias naturales, ciudadanía, plástica, teatro)?
Pregunta de investigación 3	¿Existen estrategias motivacionales L2 en las clases impartidas por profesores con formación en AfL que no están identificadas en clases no-AfL?
Pregunta de investigación 4	¿Existe una relación entre el uso, por los profesores, de técnicas de AfL y las estrategias motivacionales L2 observadas en las clases de AICLE?
Parte 2 **(Capítulo 7)**	**Los efectos de la evaluación en la motivación de los alumnos y los sentimientos durante la lección**
Objetivo principal: *Determinar los efectos de la evaluación en la motivación auto reportada de los alumnos y sus sentimientos en diferentes situaciones en las clases*	
Pregunta de investigación 5	¿Se puede detectar una relación entre el tipo de la evaluación utilizado por profesores y la motivación auto reportada de los alumnos? ¿Qué grupo demuestra más motivación: el grupo AfL o el grupo no-AfL?

Pregunta de investigación 6	¿Cómo describen alumnos AfL y no-AfL sus sentimientos en determinadas situaciones en el aula?
Parte 3 **(Capítulo 8)**	**La reflexión de alumnos de bajo rendimiento en su propio aprendizaje**
Objetivo principal: *Medir la forma en que alumnos de bajo rendimiento reflexionan sobre su propio aprendizaje con el uso de varios recursos lingüísticos.*	
Pregunta de investigación 7	¿Cómo reflexionan alumnos de AfL y no-AfL de bajo rendimiento sobre su propio aprendizaje y su experiencia en el aula?

Este estudio está dividido en tres partes que reflejan los tres enfoques, cada una con sus preguntas de investigación. El objetivo del primer enfoque es identificar estrategias motivacionales L2 que se encuentran en un corpus de clases AfL y no-AfL. El segundo enfoque está centrado en el punto de vista de los alumnos. Los datos fueron obtenidos a través de un cuestionario entregado a una muestra reducida de participantes: alumnos de la clases de ciudadanía (N=40). El cuestionario midió la motivación de los alumnos, y sus sentimientos en ciertas situaciones de la clase. El tercer enfoque también se centra en los alumnos, y analiza la forma en que los alumnos autoevalúan y hacen comentarios sobre su experiencia en el aula. Los datos fueron obtenidos de entrevistas grabadas con alumnos de bajo rendimiento (N=6).

La siguiente sección presenta un resumen estructural de la tesis.

Estructura de la tesis

La primera parte de la tesis (capítulos dos, tres y cuatro) presenta una visión global del marco teórico. El capitulo segundo expone el concepto de la evaluación, presentando varias formas de evaluación y su impacto en el aprendizaje de los alumnos. Detalla también el surgimiento de la evaluación para el aprendizaje y su uso práctico en el aula. El capítulo tercero repasa varias teorías de la motivación para aprender una segunda lengua y el desarrollo del *Motivated Orientation of Language Teaching* (MOLT): un esquema que ayuda en la observación e identificación de estrategias motivacionales L2 presentes en el aula. En el capítulo

cuarto introduce *Appraisal Theory,* una extensión de la metafunción interpersonal en la Lingüística Sistémico-Funcional (LSF) que representa un marco para el análisis en esta sección del presente estudio.

El capítulo quinto describe la metodología del estudio, aportando las preguntas de la investigación, participantes y método usado para recopilar el corpus. En el capítulo se incluye además una descripción de la selección de datos, instrumentos usados y procedimientos de análisis para cada uno de los tres enfoques.

Los capítulos sexto, séptimo y octavo se dedican a presentar los resultados. El capítulo sexto analiza las estrategias motivacionales L2 utilizados por los profesores AfL y no-AfL con respeto a la duración, frecuencia y distribución de las mismas. Para mostrar la diferencia en el uso de las estrategias, incluyen ejemplos analizados de la interacción en el aula. El capítulo séptimo introduce el punto de vista de los alumnos, presentando datos obtenidos de una muestra de alumnos de las clases de ciudadanía. Se presentan los resultados obtenidos de un cuestionario que mide la motivación y sentimientos de los alumnos en varias situaciones en el aula. En el capítulo octavo, se analizan los datos obtenidos de las entrevistas con una muestra de alumnos de bajo rendimiento para determinar la forma en que estos alumnos reflexionan sobre su propio aprendizaje y experiencia en el aula.

La tesis sigue con una discusión (capítulo noveno) de los resultados. Finalmente, en las conclusiones (capítulo décimo) se realiza un resumen el estudio y se presentan las limitaciones y propuestas para futuras investigaciones.

Resumen del capítulo

La introducción comenzó con el objeto, el propósito y el alcance del presente estudio. Explicó la necesidad de investigación en el área de la evaluación y la motivación en el campo de AICLE. También destacó la necesidad de medir los efectos de la evaluación para el aprendizaje en alumnos que están aprendiendo inglés como lengua extranjera. El capítulo proporcionó un resumen de la implantación de AICLE como un modelo de la educación bilingüe en Europa, describiendo su desarrollo en España en las últimas décadas. También se

presentaron los objetivos y preguntas de investigación del estudio. La introducción concluyó con un resumen de la estructura de la tesis. Los siguientes capítulos se dedican a dar un marco teórico de la literatura relevante al contexto de la presente investigación.

Conclusiones

Introducción

En el capítulo 9 se examinaron los resultados obtenidos en los tres enfoques de la tesis, que incluyó: un análisis de unidades didácticas de AfL y no-AfL para identificar las estrategias L2 de motivación; la motivación general y los sentimientos de los alumnos durante las clases; y, finalmente, el punto de vista de los alumnos de bajo rendimiento. Este capítulo presenta las conclusiones, comenzando con un resumen del estudio y una presentación de las conclusiones obtenidas de las tres perspectivas analizadas. A continuación, se presentan las sugerencias para profesores AICLE y las limitaciones del estudio. El capítulo cierra con propuestas para futuras investigaciones y comentarios finales.

Resumen del estudio

Esta tesis estudió aspectos de la evaluación en AICLE, teniendo en cuenta los efectos de la implementación de AfL y su efecto en el discurso motivacional en la clase y la motivación de los alumnos. El impulso del estudio es la necesidad de investigar la evaluación en el modelo AICLE, que ha sido destacado como un área emergente de investigación (Llinares, Morton and Whittaker, 2012; Coyle, 2010; Maggi, 2012; Barbero, 2012). Actualmente, el papel de la evaluación en AICLE está siendo investigado por el proyecto AECLIL, financiado por la Comisión Europea. El objetivo de este proyecto es revisar estrategias de evaluación que están impartidas por profesores AICLE, y proponer nuevas estrategias que puedan ser beneficiosas para los alumnos. Sin embargo, se necesita más investigación empírica para determinar los efectos de estas estrategias para alumnos que están aprendiendo inglés como lengua extranjera (Heritage et al., 2013; Alvarez et al., 2014) y para alumnos de bajo rendimiento, lo que ha impulsado la investigación llevada a cabo en esta tesis.

El capítulo 2 revisó las publicaciones relevantes sobre la evaluación, describiendo los retos y problemas con la evaluación tradicional. También, se elaboró el propósito de la AfL para complementar la evaluación sumativa para identificar lagunas de conocimiento. Los capítulos 3 y 4 proporcionaron un análisis bibliográfico de la

motivación y *Appraisal Theory*. También describieron los esquemas usados en el análisis de los datos. El capítulo 5 presentó la metodología del estudio, que tenía el objetivo de determinar la relación entre la AfL y las estrategias motivacionales L2 usadas por los profesores durante la clase Este objetivo se logró mediante un análisis de un corpus, identificando las estrategias motivacionales L2 en clases AfL y no-AfL (Capítulo 6). El segundo objetivo fue examinar la motivación y sentimientos de alumnos AfL y no-AfL (Capítulo 7). Estos resultados se compararon con aquellos de la primera parte del estudio para intentar determinar la relación entre las estrategias identificadas en clases de AfL y no-AfL y la motivación de los alumnos. Finalmente, el objetivo del tercer enfoque (Capítulo 8) era investigar la manera en que alumnos de bajo rendimiento evaluaron su propio aprendizaje y experiencia en el aula.

Dentro del limitado alcance del estudio, se observó un aumento de estrategias motivacionales L2 en las clases de AfL. Esto resultó en un discurso más motivacional. La frecuencia y la duración de las estrategias fueron más altas en las clases de AfL. Además, se encontró una mayor variedad de estrategias en las unidades AfL, en especial en las clases de ciencias y teatro. Un análisis de las unidades AfL mostró que los profesores prestaban más atención a la hora de clarificar los objetivos y criterios del aprendizaje, haciendo referencia a los mismos desde el inicio de las clases y durante las mismas. Además, el uso de *feedback,* o retroalimentación, en clases AfL era más positivo e incluyó propuestas específicas de mejora, mientras que en las clases no-AfL, el *feedback* era más neutro. *Process feedback,* o retroalimentación sobre el proceso de aprendizaje, era la única estrategia motivacional L2 encontrada solo en las clases AfL. La autoevaluación y evaluación por pares aparecieron frecuentemente en las clases de AfL y muy poco en las clases no-AfL. A través de la autoevaluación y evaluación por pares, los alumnos AfL participaron en intercambios para evaluar su progreso en áreas de contenido y lenguaje basado en los criterios establecidos por los profesores. El trabajo en grupos y parejas fue más frecuente en clases de AfL, fomentando que los alumnos se involucrasen en debates sobre el tema de la unidad. A través de estas

sesiones, los alumnos participaron de manera activa en el aula. Finalmente, la profesora de ciencias AfL usó técnicas para dar un papel más activo a los alumnos en el proceso de aprendizaje. Esto llevó a un aumento de estrategias motivacionales, tales como promoción de la autonomía, creando curiosidad y aumentando la atención de los alumnos.

El segundo enfoque del estudio midió la motivación de los alumnos AfL y no-AfL. A pesar de una frecuencia más alta de estrategias motivacionales L2 en clases de AfL, no se detectaron diferencias significativas entre la motivación de los alumnos AfL y no-AfL en el cuestionario. Una posible explicación es que la muestra de alumnos que rellenaron el cuestionario estuvieron expuestos a un conjunto de estrategias motivacionales similares en sus clases de ciudadanía, lo que propició que la motivación de ambos grupos fuese parecida. Aun así, los alumnos AfL expresaron una mayor incertidumbre, con respecto a sus propias capacidades lingüísticas, que los alumnos no-AfL. Este resultado también apareció en las entrevistas con alumnos AfL de bajo rendimiento, observándose que dichos alumnos eran más críticos acerca de su capacidad de aprendizaje en una segundo lengua y sus habilidades de hablar inglés correctamente en el futuro.

Una posible explicación de estos resultados, considerando la exposición de los estudiantes a su grupo de iguales y su autoevaluación, podría ser que los alumnos AfL suelan ser más preparados para poder evaluar sus habilidades o capacidades. Sin embargo, los alumnos AfL expresaron sentimientos positivos en ciertas situaciones en el aula. Estas situaciones incluyeron aquellas que pueden provocar ansiedad, por ejemplo, cuando levantan la mano, cuando el profesor les pide contestar o cuando otro compañero les ayuda. Otros resultados indicaron que los alumnos AfL mostraron menos ansiedad relacionada con los exámenes, lo que puede ser consecuencia de la dinámica de las clases de AfL que dan menos importancia a estos exámenes. Finalmente, todos los alumnos de bajo rendimiento en clases AfL y no-AfL expresarón una valoración positiva relacionada con el aprendizaje del inglés.

Los resultados de este estudio contribuyen a un entendimiento acerca de cómo las técnicas de AfL pueden integrar estrategias motivacionales de una forma más sistemática en el contexto de AICLE. Este es uno de los primero estudios que ha examinado los efectos de AfL en AICLE empíricamente con el uso de un corpus, es de esperar que se realicen más estudios de este tipo en el futuro.

Para entender las técnicas de AfL y sus efectos en la motivación, es necesario ver su aplicación en el aula. Esta fue la intención de Black y Wiliam al principio de su investigación de la evaluación formativa. Al comienzo, los dos investigadores compararon el aula a una "caja negra" con una variedad de estímulos entrando (alumnos, profesores, normas) y saliendo (resultados de exámenes, alumnos con mejor informados, profesores satisfechos) (Black y Wiliam, 1998b). El problema que identificaron Black y Wiliam es una falta de conocimiento de lo que estaba pasando dentro de la "caja negra". La base de su investigación era desarrollar un imagen más clara del papel de la evaluación formativa en el aula a través de un análisis biográfico exhaustivo. En el caso de esta investigación, mi esperanza es que, a través del análisis extensivo del corpus (que incluyó seis unidades didácticas con un total de 14 clases) se haya dado una visión más completa de la "caja negra". La meta del estudio era llegar a entender las técnicas, las interacciones y los procesos que forman parte de la AfL.

Este estudio también destaca la necesidad de tener en cuenta los efectos de la AfL en los alumnos, puesto que hasta el momento las investigaciones previas de la AfL no han tenido en cuenta el punto de vista de los estudiantes (Flórez y Sammons, 2013). Este hecho se hace especialmente relevante en el caso de los alumnos de bajo rendimiento, uno de los grupos más afectados por AICLE, debido a las demandas del contenido y el lenguaje. La necesidad de investigación centrada en este grupo ha sido destacada por diversos autores (Dobson, Pérez and Johnstone, 2010), proponiéndose como uno de los objetivos de esta tesis el investigar las capacidades de los alumnos de bajo rendimiento en la autoevaluación de su aprendizaje. Los cambios frecuentes en la política educativa ocurren sin considerar sus efectos para los alumnos, lo que puede llevar a que los éstos puedan ser negativos, por ejemplo,

cambios políticos que plantean exámenes externos de gran repercusión para los alumnos (Stiggins, 2007, Black et al. 2002, Shohamy, 2001). No hay que perder de vista nunca el hecho de que los alumnos son los beneficiarios de estas técnicas de evaluación, y es muy importante entender los efectos relativos al proceso educativo desde su punto de vista.

Implicaciones pedagógicas para profesores AICLE

Una de los objetivos de los análisis llevados a cabo en esta tesis era reflexionar sobre las implicaciones para la práctica docente, correctamente sobre el uso de la AfL para hacer un discurso más motivacional en el aula AICLE. El análisis el corpus, reveló que un gran numero de las técnicas de AfL tenían el mismo propósito que las estrategias motivacionales L2. Por ejemplo, la importancia de **aclarar los objetivos de aprendizaje** ha sido destacado en AfL (Black and Wiliam, 1998a), lo que implica el desarrollo de varias técnicas AfL para integrar los objetivos de aprendizaje. En este sentido, el uso de carteles WALT (*We are learning to…*) y WILF (*What I'm looking for…)* ocurrió al principio de casi todas las clases de AfL, y se corresponde con la estrategia motivacional L2 *signposting*. A través de aclarar los objetivos constantemente y ofrecer *feedback* a los alumnos sobre su progreso, el profesor puede ayudar a los alumnos hacer metas de aprendizaje. El uso de frases como "*I can"* también permite identificar las metas de aprendizaje y ayudar a los alumnos en la evaluación de sus capacidades.

La integración de la **autoevaluación y evaluación en parejas** sirve para desarrollar las capacidades metacognitivas de los alumnos y hacerlos más activos en el aula. Esta técnica también corresponde a la categoría *elicitation of peer and self-correction* en el esquema MOLT. La evaluación entre pares da un papel más activo a los alumnos en el proceso de reflexión sobre su propio aprendizaje (Sadler, 1989), contribuyendo a formar alumnos más críticos, ya que les entrena reflexionar en base a criterios establecidos. Estas sesiones dan un momento de reflexión en las que los alumnos pueden considerar si los objetivos de la lección se han alcanzados. El uso de técnicas AfL se ha mostrado como una forma de involucrar a los alumnos en las

clases y proporciona a los profesores la posibilidad de facilitarles información especifica acerca de sus logros, ayudándoles así a superar los objetivos.

Este estudio también destaca la necesidad de **poner a los alumnos en un papel más activo junto con el profesor, actuando éste último como un mediador que facilita el apoyo necesario.** Dar autonomía a los alumnos implica un avance hacia la responsabilidad de su propio aprendizaje. Esta autonomía ayuda a los alumnos a ser más independientes, a pensar en en los objetivos y alcanzarlos con el apoyo de sus compañeros y con el profesor. A continuación de estas sesiones, debería de promoverse una discusión de los resultados relevantes, dando la oportunidad a los alumnos para compartir información y evaluar a sus compañeros. Esta técnica puede ser beneficiosa para el establecimiento de objetivos de contenido y lenguaje, trabajando con los alumnos para superar estas metas. Todas estas técnicas comparten el propósito de asegurar que los alumnos sean más responsables con su propio aprendizaje, mostrándoles los objetivos a conseguir y ofreciéndoles una guía para obtener dicha meta.

Limitaciones del estudio

Este estudio está basado en una muestra pequeña de profesores (N=4) y alumnos (N=132). Aunque un corpus mucho más amplio estaba recopilado y transcrito para el estudio, se seleccionaron las seis unidades didácticas (14 clases /71,504 palabras) aquí presentadas, puesto que permitían realizar una comparación más detallada por asignaturas afines. El análisis manual temporal ha facilitado una imagen más específica de las técnicas AfL en el contexto de las clases. Este tipo de análisis tan exhaustivo no hubiera sido posible con un corpus más amplio. Aunque el uso de grupos de muestras pequeñas es común en investigación AfL (Flórez and Sammons, 2013), estudios de gran escala son necesarios para evitar la generalización de los efectos de AfL basados en muestras limitadas (Black and Wiliam, 2003). Además, el contexto de investigación en este estudio incluyó sólo colegios bilingües en Madrid, España, donde las técnicas AfL habían sido implementados recientemente. La

medida de los efectos de estas técnicas requerirá de una mayor investigación en el futuro, habida cuenta amplia difusión que se espera de la AfL.

Investigación adicional

Aunque el papel de la evaluación está siendo investigado actualmente a través del proyecto AECLIL (Barbero, 2012), existe la necesidad de explorar las posibilidades de la AfL y la evaluación formativa para alcanzar metas de contenido y lenguaje. Este ha sido identificado como uno de los retos más importantes en AICLE, siendo la evaluación formativa considerada como una posible solución para conectar las dos áreas (Coyle, 2010). Sin embargo, la implementación de AfL ha sido limitada en los contextos AICLE, existiendo una necesidad de mayor investigación para determinar los beneficios para profesores y alumnos. Tal investigación podría incluir el uso de técnicas AfL para identificar lagunas de contenido y lenguaje, y sugerir cómo estas técnicas pueden influir en la experiencia de aprendizaje de los alumnos. Así mismo, la evaluación de estas nuevas técnicas se puede realizar a través de entrevistas con profesores AICLE que hayan tenido formación AfL para incluir el punto de vista de los participantes. También sería interesante la realización de estudios longitudinales para observar las técnicas de AfL usadas por los profesores y sus efectos en el aprendizaje y motivación de los alumnos. No obstante, el propósito de esta investigación era realizar un análisis profundo de las grabaciones para descubrir la relación entre la AfL y las estrategias de motivación, así como para entender las reacciones de los alumnos.

Comentarios finales

La introducción del Aprendizaje Integrado de Contenidos y Lenguas Extranjeras ha tenido un éxito considerable como una iniciativa bilingüe en Europa, pero aun así es un modelo nuevo todavía en proceso de desarrollo (Dalton-Puffer, 2008). La evaluación en AICLE es una de las líneas de investigación emergentes que busca ofrecer una solución a las necesidades de aprendizaje integrado de contenido y lengua para los alumnos. Debido a los retos que suponen estudiar en una lengua extranjera, es posible que los alumnos de AICLE requieran un nivel de motivación

más alto que los estudiantes de que estudian en su primera lengua. La motivación es un concepto complejo y abstracto, y precisamente por eso, una de las metas comunes de los profesores como investigadores es buscar maneras para fomentar y mantener la motivación de los alumnos. La autora espera que este estudio haya servido para llegar a un mayor conocimiento de cómo la AfL puede contribuir a este objetivo en contextos AICLE. Además, se espera, que esta tesis pueda dar lugar a más investigación empírica sobre el uso de la evaluación en AICLE para promover la motivación en los alumnos.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

"Assessment for learning is a gift we give our students. It is a mirror we hold up to show them how far they have come. It is a promise that we will use assessment not to punish or reward, but to guide them on their learning journey."

-Jan Chappuis, Assessment Training Institute

1.1 Background and motivation for the research

This study focuses on assessment for learning (AfL) and its relationship to motivation in the CLIL context. The aim of AfL is to use assessment as a tool to help learners identify and fill learning gaps with mediation from the teacher, as defined by the Broadfoot et al. (2002):

> (AfL is) The process of seeking and interpreting evidence for use by learners and their teachers to decide where learners are in their learning, where they need to go and how best to get there (Broadfoot et al., 2002: 2-3).

To date, a significant amount of literature has focused on the benefits of formative assessment, or assessment for learning. However, limited research has explored its effects on English language learners (ELLs) (Alvarez et al., 2014). The need for research in this area has been cited as necessary by Heritage et al. (2013) to determine the potential of assessment in meeting content and language needs.

> As use of formative assessment expands, it is imperative to examine the potential impact of formative assessment practices on both the academic achievement and language learning of ELL students and to explore how formative assessment practices could be tailored to meet the specific needs of these students (2013).

This study aims to address this gap in the literature, investigating AfL in the context of Content and Language Integrated Learning (CLIL) classrooms, in which ELLs

learn language and content simultaneously. CLIL refers to any "activity in which a foreign language is used as a tool in the learning of a non-language subject in which both language and subject have a joint role" (Marsh 2002:58). That is, situations in which content is taught through the medium of a foreign language, with both given equal consideration. CLIL is a teaching approach introduced in the 1990's that has become integrated into educational settings throughout Europe and is discussed further in section 1.3. Despite its widespread implementation over the past 25 years, the role of assessment in CLIL has been cited as an under researched area needing attention (Llinares, Morton and Whittaker, 2012; Coyle, 2010; Maggi, 2012; Barbero, 2012).

Research on assessment in CLIL has mainly focused on the role of language (Höning, 2009) or assessments of language proficiency (Wewer, 2014). The Assessment and Evaluation in CLIL (AECLIL) project addresses this gap in research (Barbaro, 2012). Funded by the European Commission (EACEA), the aim of this project is to develop tools to be used in CLIL assessment. One of the main concerns regarding CLIL has been the fear of not meeting objectives due to the added teaching and learning load resulting from teaching content in a foreign language (Clegg, 2012). Formative assessment has been identified as a means of addressing content and language goals (Kiely, 2009; Coyle, 2010) and is now being implemented in certain CLIL classrooms in Spain, the context in which this study has been carried out.

Due to pressure on students to master language and content related objectives, the need for motivational teaching practice in CLIL is exceptionally high. This is especially true in the case of lower achieving learners. The focus of CLIL has traditionally been on higher achieving students (Denman, Tanner and De Graff, 2013). The need for investigation on lower achieving CLIL students is cited as an area needing further attention (Dobson, Pérez and Johnstone, 2010). Research on motivation in CLIL classrooms has focused on secondary schools, with few studies focusing on primary school settings (Lasagabaster and López Beloqui, 2015). This study seeks to determine the potential of AfL as an effective way for CLIL teachers to

promote a more motivational classroom environment for young learners and address the needs of lower achieving students.

1.2 Purpose and scope of the study

This study examines the relationship of assessment for learning (AfL) and motivation in CLIL classrooms using three perspectives. The first analyzes a classroom corpus comparing L2 motivational strategies put into practice by AfL and non-AfL teachers. The second focuses on the students' point of view by analyzing self-reported motivation of CLIL learners in AfL and non-AfL settings. This perspective also measures the feelings of these students in different classroom situations. The third perspective aims to determine the ways in which lower achieving students are able to assess themselves and their learning environment through interviews. Research shows the positive effects of AfL on student motivation (Ross, Tronson and Siegenenthalter, 2006; Butler, 1988; Schunk, 1996), mostly based on questionnaires, interviews and focus groups. This study employs a mixed methods approach to provide a fuller picture and deeper understanding of the data (Johnson, Onwuegbuzie & Turner, 2007) by combining quantitative and qualitative analyses and multiple data sets.

To date, the majority of AfL studies have taken place within the United Kingdom and United States, two regions in which formative assessment is widely used and ingrained in teacher training. The setting for this study takes place in Madrid, Spain, a rich language-learning environment that has been expanding efforts over the past two decades to provide language-learning opportunities to learners starting at an early age. The study is conducted in schools belonging to bilingual programs in which CLIL is implemented in 40% of the curriculum. The following sections provide an overview of CLIL and its implementation in Spain, contextualizing the setting for the present study.

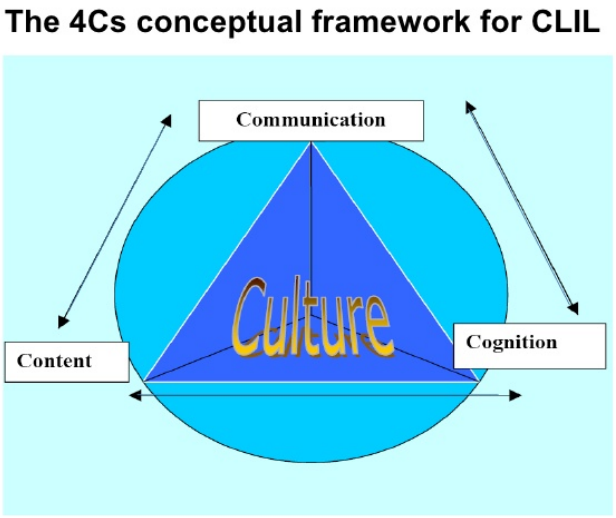
1.3 Content and Language Integrated Learning (CLIL)

The push for an increasingly multilingual European community has been an ongoing process over the past two decades. The European language preservation act began in the mid-1990's, with the publication of the *White Paper on Education and Training* by the Commission of European Communities calling for a serious commitment to bilingual education starting as early as nursery school (Commission of European Communities, 1995). Since the English language has evolved into a lingua franca both within the European community and beyond, the trend toward adopting English as a medium of instruction began to emerge from 1995 to 2005 (Marsh 2006).

Since its initial appearance in the 1990's, the influence of CLIL has been widespread. However, according to the Eurydice report on CLIL in 2006, the model is not fully complete and much needs to be studied to develop its theoretical underpinnings (Dalton-Puffer, 2008). Since this realization, scholarly investigation into CLIL has been rapidly expanding throughout Europe as researchers seek to provide a theoretical basis and explain the advantages and limitations of its implementation. Through an extensive number of research studies on CLIL conducted within the European community, researchers now have an understanding of how the approach affects language learning.

In addition to helping prepare students for internationalization by improving learners' intercultural communication competence, CLIL has been shown to increase motivation to learn foreign languages, improve vocabulary and enhance incidental learning through a focus on meaning and communication (Lasagabaster, 2008). The 4C's framework (Coyle, 1999) provides the pillars of CLIL teaching: content, communication, cognition and culture.

Figure 1.1: The 4C's of CLIL



Coyle (1999, 2005)

The greatest benefit of CLIL, which distinguishes it from traditional English as a Foreign Language (EFL) learning, is the focus on oral communication between teachers and learners, as well as the learners themselves, which is shown to enhance language development (Lasagabaster, 2008). Through this increased focus on communication, a variety of language competencies are affected, some of which are illustrated in table 1.1.

Table 1.1 Language competencies favorably affected or unaffected by CLIL

Favorably affected	**Unaffected or Indefinite**
Receptive Skills	Syntax
Vocabulary	Writing
Morphology	Informal/ non-technical language
Creativity, risk-taking, fluency, quantity	Pronunciation
Emotive/ affective outcomes	Pragmatics

(Dalton-Puffer, 2008: 5)

CLIL improves students' listening skills, vocabulary (namely subject- specific), fluency and ability to express themselves in the foreign language. Moreover,

research shows that participating in CLIL programs has a positive influence on student motivation (Seikkula-Leino, 2007, Lorenzo et al., 2009; Lasagabaster, 2011; Doiz, Lasagabaster and Sierra, 2014; Sylvén and Thompson, 2015).

Despite a high volume of research regarding the benefits, criticism of CLIL cites challenges such as: isolation amongst CLIL teachers (Cross and Gearon, 2013); selection in bilingual programs which may be exclusionary or create a hierarchy among non-CLIL students (Bruton, 2011); and striking a balance between language and content (Rowe and Coonan 2008). The limited guidelines for implementing CLIL teaching practice are cited as an issue, considering the multi-dimensional nature of addressing language and content (Meyer, 2010). As mentioned previously, this issue is particularly relevant in CLIL assessment (Coyle, 2010), which the development of the AECLIL project (Barbaro, 2012) seeks to address. Tools developed include: a variety of rubrics meant to measure students' general performance and communicative language skills (Barbero, 2012); grids which act as a tool to assess student's essays and oral performances (Barkovsa, 2012); and finally, peer and self-evaluation rubrics (Maggi, 2012). While the development of such tools is necessary to define the role of assessment in CLIL, research is also needed to observe the implementation of assessment practice in CLIL classes and determine the effects of such assessment on learners.

1.4 Bilingual education in Spain

In the past several decades, bilingual CLIL-style education has been implemented at a progressively increasing rate within Spanish Autonomous Communities. According to a 2006 study by the European Commission, Spain had fallen behind the language learning curve in the European Union with only 17% of Spanish citizens claiming to be able to communicate effectively in a foreign language, a number far below the 56% average of many European citizens. In fact, 56% of Spaniards admitted to being completely monolingual, falling into the lowest tier of language competency in European Member states with Italy, the United Kingdom and Ireland (European Commission, 2006). Thus, CLIL programs have risen in Spain over the

past decade because "they are believed to represent the best way to augment the traditionally low foreign language command among Spanish students" (Ruiz de Zarobe and Lasagabaster, 2010).

To try to improve foreign language competence, Spain began launching language learning initiatives in the 1990's focusing on English, which is cited as the most commonly adopted language in Europe (Eurydice, 2006). In 1990, the *Ley Orgánica de Ordenación General del Sistema Educativo* (an educational reform act) was passed, stipulating the introduction of a first foreign language by the age of eight (Gobierno de España, Ministerio de la Presidencia, 1990). Since then, the starting age has dropped to age three, when children can begin their first year of nursery school. Reasoning for this change is based on the fact that the younger the learner, the more receptive they will be to foreign language learning (Lasagabaster, 2008; Halbach, 2009).

As a result of these changes in educational policy with an emphasis on multilingualism, the British Council in conjunction with the Ministry of Education and Science launched the Bilingual Project in 1996, implementing CLIL in schools throughout Spain. The project began in primary state schools with the goal of expanding to secondary schools by 2004, and has been successful in promoting its educational purpose and preparing the younger generation to enter the increasingly multilingual European workforce in the future (Reilly and Medrano, 2009). Since its inception, the program has continued expanding throughout Spain and is well established in Madrid. An external evaluation of the British Council Bilingual Project concluded that the program is meeting language related goals, noting that students have a "fluent and confident command of English, including technical vocabulary and the production of extended utterances" (Dobson, Pérez and Johnstone, 2010). Subsequently, in 2004, the *Comunidad de Madrid*[1] (CAM) began a similar bilingual project in public primary and secondary schools in Madrid and the surrounding

[1] The *Comunidad de Madrid* refers to the regional governing body of the Madrid Autonomous Community.

areas. The program has been expanding over the past ten years, with over 300 primary schools and nearly 100 secondary schools following the bilingual program to date (Comunidad de Madrid website). Thus, the bilingual educational system in Madrid reflects a growing effort to promote English language learning in young learners with over 40% of academic subjects taught in English.

1.5 Research Background

The impetus for the present study stems from contact between the UAM-CLIL research team and the British Council in 2009. The team requested access to various primary bilingual schools for the purposes of conducting a study on assessment practices. In response to this request, the British Council informed the team of a new initiative of bringing assessment for learning into the bilingual program in Madrid. The focus of the study was then placed specifically on AfL and its impact on learners in the program. At the time, only a small number of teachers had received AfL training. Therefore, the goal of the research was to determine how these techniques were influencing students' learning and whether more widespread training would be beneficial. After a review of relevant AfL literature, I decided to focus this research on the relationship between AfL and motivation, an area in which empirical research was needed.

1.6 Participants and research context

Participants for the study came from four primary schools in the Comunidad de Madrid belonging to the two bilingual programs mentioned previously. Four teachers were involved in the study: two worked in British Council bilingual schools and had undergone training in assessment for learning techniques; and two worked in CAM bilingual schools and had no exposure to AfL training. The subjects taught included: science (natural and social), citizenship, art, drama and English, though for the purposes of this study, only content area subjects were included. The students (10-12 years old) were in 5th and 6th grade of primary education, and had been participating in the bilingual program since the age of three. The four schools were

located in various suburban areas of the Madrid Autonomous Community, ranging from middle to upper-middle class.

1.7 Objectives and research questions

The present investigation employs a variety of approaches to determine the relationship of AfL and motivation. The data is drawn from a classroom corpus collected throughout the 2010/2011 academic year in a variety of subjects: science, citizenship, art and drama. Additional data comes from questionnaires and interviews given to a small sample of students.

The structure of the study, objectives and research questions are the following:

Part 1 **(Chapter 6)**	**Motivational L2 strategies in AfL and Non-AfL lessons**
Main objective: To compare discourse in AfL and Non-AfL classrooms with a focus on second language learning motivational strategies.	
RQ1	Do the frequency and distribution of second language motivational strategies differ depending on the use of AfL?
RQ2	How does the duration of these L2 motivational strategies vary depending on the subject (science, citizenship, art, drama)?
RQ3	Are there any L2 motivational strategies found in AfL lessons that are not identified in non-AfL lessons?
RQ4	Is there a relationship between teachers' use of AfL techniques and L2 motivation strategies observed during CLIL lessons?
Part 2 **(Chapter 7)**	**The effects of assessment on student self-reported motivation and feelings during the lesson**
Main objective: *To determine the effects of assessment techniques on student self-reported motivation and their feelings in certain classroom situations.*	
RQ5	Can any relation be seen between the type of assessment used by teachers and student's self-reported motivation? Are students in AfL classes more or less motivated than their non-AfL peers?

RQ6	How do AfL and non-AfL students describe their feelings in the context of certain classroom situations?
Part 3 **(Chapter 8)**	**Lower achieving students' reflections on their own learning**
Main objective: *To determine the extent to which lower achieving learners express metacognitive reflection on their own learning through the use of various linguistic resources.*	
RQ7	How do lower achieving AfL and non-AfL students reflect on their own learning and classroom environment?

The study is divided into three parts with research questions for each. The first part seeks to identify second language (L2) motivational strategies present in AfL and non-AfL classroom interactions using a corpus of recorded lessons. The second part is student-centered, grounded in data obtained from a questionnaire given to a sample of participants from citizenship lessons (N=40) measuring self-reported motivation and feelings in the class. The third part, also student-centered, analyzes the ways in which students self-evaluate and comment on their learning experience. Data from this part of the study comes from recorded interviews with select lower achieving students (N=6).

The following section provides a more complete structural overview of the dissertation as a whole.

1.8 Structural overview

The first part of this dissertation (Chapters 2 to 4) gives an overview of the theoretical background. Chapter 2 discusses the concept of assessment, reviewing various forms of assessment and their impact on student learning before detailing the emergence of assessment for learning and its practical use in the classroom. Chapter 3 reviews several theories of L2 motivation and the development of the Motivated Orientation of Language Teaching (MOLT) scheme, a tool for observing and tracking L2 motivational strategies present in the classroom. Chapter 4

continues by introducing Appraisal Theory as an extension of the interpersonal metafunction in Systemic Functional Linguistics (SFL), which is used as a framework for analysis in the present study.

Chapter 5 outlines the methodological approach of the study, describing the research questions, participants and method of collecting the corpus. This is followed by a description of the data selected, instruments and analysis procedure for each of the three perspectives of the study.

Chapters 6-8 are dedicated to the presentation of the results. Chapter 6 analyzes the motivational L2 strategies used by AfL and non-AfL teachers focusing on the differences in duration, frequency and distribution. Classroom extracts for each strategy are given to show the differences in usage. Chapter 7 explores the student perspective based on a sample group of learners from citizenship classes. This chapter presents the results obtained from a questionnaire measuring students' self-reported motivation and feelings in certain classroom situations. Chapter 8 analyzes data from interviews with lower achieving learners to determine how these students reflect on their own learning and their learning environment.

The dissertation continues with a discussion (Chapter 9) of the relevant findings. The conclusions are found in Chapter 10, which summarizes the study as a whole, addressing limitations and providing suggestions for further research as well as implications for teachers.

1.9 Chapter summary

The introduction began with the motivation for the present dissertation, followed by the purpose and scope of the study. The need for research in the area of CLIL assessment and motivation was emphasized, as well as the effects of assessment for learning on English Language learners. This was followed by an overview of CLIL as a model for bilingual education in Europe, noting its emergence in Spain over the past several decades. The objectives of the study were then described, providing aims and research questions to be addressed. The introduction concluded with a

detailed structural overview of the dissertation. The next several chapters provide theoretical background of the relevant literature, setting up the context for the present investigation.

Chapter 2: Theoretical framework: Assessment and AfL

2.1 Introduction

Learning and assessment have been co-evolving over the years (Cumming and Maxwell, 1999). While widespread use of assessment based on standardized testing in order to raise standards in schools continues to be the driving force behind many changes in education (Harlen and Crick, 2002), in recent years this philosophy has been changing. Still, an increased amount of importance is being placed on evaluating students' achievement based on test performance, mainly in order to provide policy makers, parents and teachers with tangible results. This trend has been especially present in the United States with the past implementation of The Elementary and Secondary Education Act (US Department of Education, 2001) and in the United Kingdom with the development of National Curriculum Testing (UK Department for Education, 2012). However, this approach can have counterproductive effects, namely the proliferation of washback, or "teaching to the test" (Rea-Dickins, 2001) while neglecting to develop the skills and learning aims intended in the curriculum (Harlen and Crick, 2002).

When considering assessment, many educators underestimate its power as a tool to develop students' knowledge rather than a simple measurement device (Black and Wiliam, 2003). Assessment for learning (AfL), or formative assessment, was developed in response to a growing need to use evaluation as a practical means of detecting and filling learning gaps (Black and Wiliam, 1998a). This chapter addresses the multifaceted nature of assessment, focusing on and making a distinction between summative and formative assessment. It discusses reforms that have taken place over the past several decades and considers the impact of AfL in the classroom.

The chapter begins with a brief introduction on assessment, exploring the ongoing dialogue between researchers, policy makers and teachers on best assessment practice. This introduction is followed by a general overview of assessment, focusing on summative assessment and assessment for learning (AfL). The effects of high

stakes testing on students are then discussed and contrasted with the implementation of AfL. Next, an overview of various AfL strategies used by teachers show how the approach is put into practice in the classroom. The chapter concludes with the discussion of several studies on AfL in relation to students' achievement, motivation and the effects on lower achievers.

2.2 Defining assessment

Assessment is the process of collecting and interpreting evidence to make judgments about students' educational achievements (Harlen, 2007), which is a central task of school administrators and teachers (Natrillo, 1987). In most academic environments, the basis for assessing students' work places the teacher or external evaluator in the position of the authority figure to judge a student's or school's progress based on evidence. There are several types of assessment, including: *diagnostic* at the beginning of an academic term to gage students' previous knowledge on a topic; *benchmark*, occurring intermittently to measure progress; *summative*, which is administered at the end of the learning process to make a judgment on students' mastery of the material, usually in the form of an exam; and *formative*, which encompasses an ongoing process that may be administered through various means during the academic year (The Eberly Center for Teacher Excellence).

The scope of evaluation expands beyond testing. Rea-Dickins (2001) addresses the question "What counts as assessment?" with three possibilities, including: tests; routine classroom teaching activity and instruction; and embedded techniques such as questioning, interaction between learner and teacher, collaboration and feedback. The goal of assessment is to determine whether or not learning has taken place, and in order for this to happen, the evaluator must look for potential clues (Black and Wiliam, 1998a). These clues may include the students' ability to focus their attention on extending a concept, using the assessment process in a different context, and demonstrate persistence, motivation and engagement to focus on a task or activity (Black and Wiliam, 1998a; Black et al., 2003).

The job of the evaluator is to interpret these clues to determine the degree of knowledge demonstrated by students and make an assessment based on this evidence. In most cases, the teacher acts as assessor, establishing tasks and criteria to measure student achievement, giving feedback and monitoring progress. However, assessment has the potential to not only measure student achievement, but also be used as a tool to enhance learning.

One of the most common forms of assessment is summative, with many teachers consistently relying on testing outcomes to evaluate students' mastery of learning objectives. Considering the amount of research in this field, this dissertation focuses on studies related to assessment for learning while providing a brief overview of summative assessment to compare the two methods and show how they might be used in tandem.

2.3 What can assessment do? What does it actually do?

The potential for assessment to improve learning and increase standards is high and has gone untapped (Stiggins, 2007). When used properly, assessment has the potential to improve student achievement (Black and Wiliam, 1998), raise standards (Black et al., 2002) and promote successful learning by urging students to take control and have ownership of their work (Assessment Reform Group, 2002) while providing tools for lifelong learning. The use of feedback acts as a means of addressing and correcting mistakes, therefore promoting further learning and increasing students' self-esteem and confidence (Stiggins, 2007). Research has indicated that, oftentimes, assessment does not live up to its full potential (Review of Secondary Education in England, 1993-97) and has the potential of provoking negative effects, especially in lower achieving students (Stiggins, 2007).

2.4 Making the distinction: Summative assessment and assessment for learning

Differences between summative and formative assessment lie not only in the type of evaluation being used, but also the function of this evaluation (Black and Wiliam, 1998a). The distinction becomes clear when considering the alternative names for

both techniques: Assessment OF learning (summative assessment) and Assessment FOR Learning (formative assessment). AfL becomes formative assessment when the teacher uses evidence gained to adapt their teaching framework or curriculum to ensure that learning needs are met (Black and Wiliam, 1998a). While both measure students' knowledge, the purposes of the measurements differ: AfL seeks to integrate assessment as a tool to increase learning (Black and Wiliam, 2003) while AoL seeks to provide a measurement of learned knowledge.

Summative assessment is concerned with the degree to which learning outcomes have been achieved (Bloom, Hastings, Madaus, 1971). Cumulative in nature, this type of assessment is administered as a means of collecting statistical data to be used as a tool for teachers to see where the students stand academically. In general, an exam is given, testing students on their knowledge gained throughout the unit, term or academic year. These examinations may be *large scale* (designed and administered by policy makers) or *small scale* (administered in a classroom setting) (Wilson 2009). Students are assessed and given a score, which is reflected in their final grade. Research has linked summative assessment to 'motivation to learn,' (Black et al. 2002) as tests and classroom grading may provide a drive for students to apply rigorous study habits to learning the material as they prepare to be tested.

According to Harmon, summative assessment is a yardstick to measure what a student can or cannot do, and furthermore is used as a measurement for the performance of schools and teachers. High stakes summative testing such as the Standardized Aptitude Test (SAT) in the United States, A-Levels in the UK or the *Prueba de Acceso a la Universidad* (PAU) in Spain have the potential to influence a student's entrance to university, and could affect employment opportunity in the future (Harmon, 1995).

Conversely, the purpose of AfL is to continuously assess students to detect where learning gaps lie and work to fill these gaps (Black and Wiliam, 1998a). AfL requires a complete change in classroom culture, creating a student-centered approach to

learning (Black et al., 2003), in which teachers deviate from the traditional role as an authority figure and instead act as a mediator (McCallum, Hargreaves & Gipps, 2005). Instead of paper-based testing, teachers encourage student output by employing effective questioning techniques (Black et al., 2003) and continually appraise student performance by offering feedback to increase competence (Tunstall and Gipps, 1996). Birenbaum et al. (2009) argue that assessment is inquiry and formative assessment is a means of interpreting evidence to see where the students are in their learning and the best way to get them where they need to be to close the learning gap (Black and Wiliam, 1998a).

Table 2.1 shows several fundamental differences between summative and formative assessment.

Table 2.1 Summative versus formative assessment

	Summative Assessment	**Formative Assessment**
	Assessment **OF** Learning	Assessment **FOR** Learning
When administered?	At the end of an academic unit, term or year	Occurs on an ongoing basis
Method of Administration	Test, quiz, exam, high stakes testing	Teacher's questions, portfolios, students set short term goals with teacher acting as a facilitator in order to achieve these goals
What is covered?	Summative review of everything that students have learned during the unit, term or academic year	Tailored to students needs; varies depending on the nature of the lesson
Type of Correction	Teacher corrects student errors	Peer correction, self correction, teacher feedback

Type of Evaluation	Numerical score, mark (A,B+, *sobresaliente, notable*[2]) or percentage.	Detailed teacher feedback (oral or written) focuses on students' strengths and weaknesses, peer feedback, self-assessment
Results given	Students are given results at a later date ranging from several days to several weeks after the exam has been completed.	Students given results instantaneously or more detailed results intermittently throughout the academic term
Goal of evaluation	To measure student achievement, certify student competence, grading	To measure student achievement, point out strengths and weaknesses, provide ways to improve and find solutions to fill learning gaps

Stiggins, 2007, Black and Wiliam 1998a

As table 2.1 shows, summative assessment takes place at the end of an academic unit and is often administered in the form of a written exam measuring students' retention of the material covered throughout the unit. The exam is returned to the student several days or weeks later with a grade (and perhaps some form of error correction), yet students are most likely not given extensive feedback on their mistakes or the opportunity to correct them. Conversely, AfL occurs at various points throughout the unit and most often on a daily basis, taking many forms, such as: portfolios, group or pair tasks, or even responses to teacher's questions. The evaluation of tasks also takes on many forms, such as feedback given by teachers (this feedback should be more than a numerical grade or error correction) and peer and self-correction. AfL also stresses the importance of explaining learning objectives to students at the beginning of a unit and setting goals. The students then

[2] Sobresaliente= Outstanding
Notable= Notable

reflect on these goals to determine whether objectives have been met and set further goals when necessary, which may involve prolonging the unit.

Over the past several decades, summative assessment has been criticized by researchers who claim that it does not offer an accurate representation of knowledge. Additionally, some claim that summative assessment, namely high stakes testing, may put undue pressure on students (Stiggins, 2007, Black et al. 2002, Shohamy, 2001). With the goal of measuring student achievement and raising standards, in many cases summative assessment may instill anxiety. Students sense the fact that they are being evaluated and become preoccupied with avoiding failure (Broadfoot et al., 1996), to the extent that learning objectives may be lost in the interest of achieving a high grade or test score.

2.5 Effects of testing on students and curriculum goals

Importance has been placed on high stakes testing in school systems worldwide. On a macro level, policy makers have become convinced that the most efficient way to measure the performance of a school system is based on standardized testing. On a micro level, teachers respond to this pressure by placing importance on the end result of learning rather than the process of learning. As a result, in many cases, classroom practice has adapted to rote learning (superficial learning of the facts as opposed to high order thinking skills) and teaching to the test (Black et al., 2003). This section addresses the consequences of such testing and the summative nature of everyday classroom assessment practice.

Pioneers of assessment for learning research Paul Black and Dylan Wiliam identify several problems with traditional assessment practice in their work.

1. Classroom assessment practices generally focus more on superficial or rote learning, concentrating on recall of isolated details, which students soon forget
2. Teachers do not review their assessments/ assignments or get peers to review them so there is little critical reflection on what is being assessed and why

3. The grading aspect of assessment is overemphasized and learning or improvement purpose of assessment is underemphasized

(Black and Wiliam 1998a)

The first issue that Black and Wiliam identify is the nature of classroom assessment, which has become focused on the memorization of material that appears on an exam, which students forget soon after the test. In many countries in Europe and North America, the performance of students on high stakes standardized tests has become the determiner of the amount of funding and resources that schools receive (National Council of Teachers of English, 2014), therefore, teachers feel pressure to emphasize this material to the exclusion of all else.

Pressure for increased test scores leads teachers to change the content of what they teach, spending a significant amount of instructional time on test preparation. Instructors transform into drill sergeants in an effort to prepare students for the exam, leaving them without time for in-depth projects and activities that develop higher cognitive skills (Harmon, 1995). However, while emphasis is placed on such test results as the final measure for a school's success, it has been shown that increased scores are unrelated to any fundamental educational improvement. Furthermore, the added pressure of such high stakes testing may lead to increased student educational failure or increased drop-out rates (Rumberger and Thomas, 2000).

The second point on Black and Wiliam's list refers to classroom assessment on a micro level, specifically the assessment criteria given to students. They note that many teachers do not reflect on the purpose of the assessment or its benefits for the learner. According to students, "good teachers" are able to provide conditions so that students can understand work, give help with difficulties and make tasks clear (Brown and McIntyre, 1993) and this also applies to assessment. A disconnect occurs when students do not understand the purpose of the assessment and are not aware of teacher expectations. Therefore, the teacher should first determine the purpose of the assessment and then make the criteria clear to students.

Finally, while summative assessment creates a sense of accountability to students (Garrison and Ehringhaus, 2007), a grade is often not enough for them to gage where their strengths and weaknesses lie and how to improve. Sometimes the grade may be accompanied by error correction, though oftentimes these corrections are returned to students at a much later time and are rendered meaningless, as they have moved on to new material. Black and Wiliam note that tests serving a summative function cannot be used to identify learner's needs (2003). Most of the time, the objective in taking a summative evaluation is simply obtaining a certain grade. As a result, some, namely underachieving students, begin to view testing as a negative experience.

In *The Power of Tests* (2001) Elana Shohamy presents qualitative data on students' reactions to testing. Many view the experience as "frightening," "unjust and unfair" and having "long term negative consequences", which Shohamy concluded have the potential to negatively impact self-esteem, with students stating, "tests always show me that I am a failure" (Shohamy, 2001). Assessment expert Rick Stiggins suggests that summative assessment divides students into groups of higher and lower achievers, or as he describes them, students who are either on a "winning or losing streak". For those on a "winning streak", excellent test scores provide a self-esteem boost and the drive for further academic success, while, for those on a "losing streak," the effect is the opposite (see Table 2.2).

Table 2.2: The assessment experience

	The Winning Streaks	**The Losing Streaks**
Results Provide	Evidence of Success	Evidence of Failure
Student Feels	Hopeful and Optimistic	Hopeless
	Empowered	Panicked then resigned
Student Thinks	"I'm doing fine"	"This hurts"
	"I succeed as usual"	"Nothing I try works"
	"I want more success"	Public failure-

		embarrassing
Student becomes likely to	Seek challenges	Seek what's easy
	Practice with gusto	Become confused about what to practice
These actions lead to	Self enhancement	Self-defeat
	Positive self-fulfilling prophecy	Negative self-fulfilling prophecy
	Curiosity, enthusiasm	Boredom, frustration, fear

(Stiggins, 2007: 24)

As the "losing streak" continues, lower achieving students may begin to feel a sense of hopelessness where anticipation of failure becomes a self-fulfilling prophecy. In many cases, test scores alone do not provide sufficient evidence to promote improvement. Even with error correction, many times students see their mistakes long after they have taken the exam and are not given the opportunity to re-take the test to correct them. Therefore, the gap grows wider as the curriculum moves on and these students become lost while their more successful peers continue to perform well (Stiggins, 2007).

Stiggins and other assessment reformers believe the only way to make the gap between lower and higher achieving students smaller is by using assessment as a tool to promote student success rather than show them their failures (Stiggins, 2004). Students' needs must be taken into account to give high quality assessment that provides a service rather than simply a measurement. Assessment for learning was developed with these goals in mind and, when used in conjunction with summative assessment, has the potential to improve student performance and learning (Stiggins, 2007). The next sections elaborate on the change in classroom culture that occurs with the integration of AfL, showing strategies used by teachers to put this type of assessment into practice.

2.6 Putting assessment for learning into practice

The aim of this section is to provide an overview of AfL techniques that may be used in the classroom, which are later referred to in the findings. The AfL trained teachers participating in this study put many of these techniques into practice. Therefore, this section identifies and later links them to strategies for motiving students in the CLIL classroom.

Assessment for learning creates new tools, requiring teachers to change classroom culture and requires radical changes in classroom pedagogy, (Black and Wiliam, 2003), calling for a more active, student centered learning approach (Ross, Tronson and Siegenthaler, 2006). Despite the findings noted in previous sections regarding the enhancement of teaching practice and student achievement through the use of assessment, most educators receive little formal training in assessment and the use of data to improve teaching and learning (Dietel, Herman and Knuth, 1991). Over the past years, the use of formative assessment and assessment for learning has become integrated into teacher training curriculums in the UK, and is encouraged by the Department of Education.

> Effective assessment is a key part of good teaching in all subjects. The best schools use regular formative assessment to assess what their pupils know and identify where they need additional support. This in turn allows meaningful feedback to individual pupils and parents. (UK Department for Education, 2013: 8)

. > The new national curriculum programs of study set out what pupils should be taught by the end of each key stage. Teachers will develop a school curriculum, which is relevant to their pupils. Schools will be able to introduce their own approaches to formative assessment. (UK Department for Education, 2013: 6).

Training in AfL techniques has not been required for public school teachers by the Spanish government. However, the British Council has provided training to teachers in Bilingual Education Project in Madrid in an effort to give teachers the training

necessary to implement AfL.

Changes in the assessment process involve several factors, including: sharing learning goals with students; providing feedback; encouraging students to self-assess based on previously established criteria; and finally, using the data gained to adapt teaching strategies. These factors are seen in the points below by Shepard (2000).

Changing classroom assessment practices involves:

1. Changing the nature of assessment conversations teachers have with students such that students develop greater knowledge and responsibility for learning goals

2. Assessing student's prior knowledge and using that information in planning better instruction to meet their needs and match their interests

3. Giving students feedback in ways that go beyond grades such that they are helped to understand what quality work or thinking looks like

4. Getting clearer about the explicit criteria for open ended/ performance tasks and involving students in self-assessing

5. Using information from students to evaluate and improve their teaching strategies

Shephard, 2000

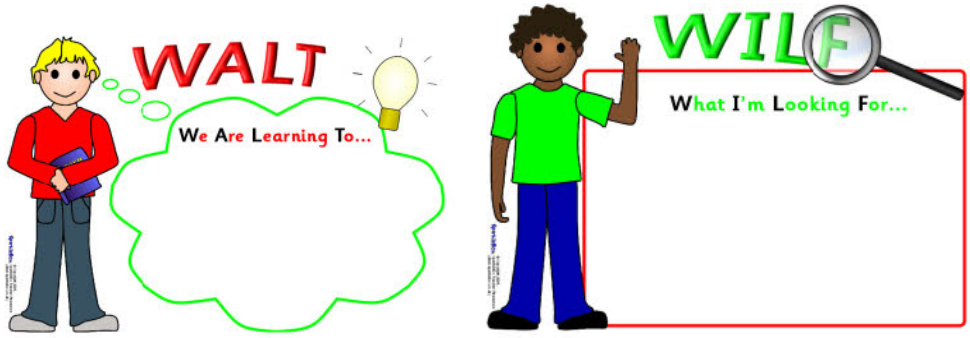
These changes are designed to help students benefit from assessment in order to facilitate learning. Therefore, five critical factors of AfL are: *stating the purpose of the lesson; effective questioning techniques; feedback; peer and self-assessment strategies and using summative assessment for formative purposes,* which are discussed in further detail in the next sections.

2.6.1 Stating the purpose of the lesson

Students who understand what they are being expected to learn and are given firm guidelines are more likely to make learning gains (Young, 2005). Therefore, an important element when facilitating student success is sharing learning goals at the onset of the lesson. Once these objectives are established, they may be shared with parents to provide support as the students work toward meeting their goals.

Putting AfL techniques into practice involves embedding "mediating artifacts" (posters, cards and other materials) into daily classroom routines (Webb and Jones, 2009). Learning objectives can be presented in a number of ways, with many AfL teachers displaying WALT and WILF posters in order to present the aims of the lesson.

Figure 2.1: WALT and WILF posters



These posters represent characters that students may become familiar with, as well as acronyms: WALT stands for **W**e **A**re **L**earning **T**o and WILF stands for **W**hat **I**'m **L**ooking **F**or. The WALT poster presents objectives of the class session or unit, which may be established by the teacher, the students or both, while the WILF poster represents the teacher's expectations. The posters may contain long or short-term aims, which the teacher shares in detail with students. These aims are then reflected upon at the end of a unit or lesson and, as a result, students are able to build learning goals and measure progress.

2.6.2 Effective questioning techniques

Rather than positioning the teacher as an all-knowing authority, the AfL teacher's role is that of a mediator, guiding the student down the appropriate learning path. The use of questions is a powerful tool for teachers to determine where students are in achieving their aims. AfL questioning techniques encompass four main phases**:** diagnostic questions, question type, wait time and interpreting student's response.

2.6.2.a Beginning the Unit: Diagnostic Questions

Asking questions should be regarded as a process of collecting evidence and may also be used as a form of diagnostic assessment to determine what students have studied in previous units (Sutton, 1995). This is an important way to measure student progress and provide continuity from course to course without repeating the same knowledge or omitting elements that students may not have previously covered. Questioning at the beginning of the unit is a key factor in measuring previous knowledge and adapting planning accordingly (Black et al., 2003). Studies show that while some teachers are adept at judging student work, many struggle to diagnose through questioning (Bennett and Kell, 1989), therefore training in this area is required.

2.6.2.b Types of questions

Question type is an important factor in implementing AfL and holds a direct correlation with engaging students in lower or higher order thinking skills. A study by Stiggins et al. (1989) investigated questioning techniques in primary classrooms over a range of subjects. The findings indicate that the most frequently used questions were recall questions, which prompted students to recount facts learned in previous lessons. However, the number of higher-order thinking questions was infrequent. When self-reflecting on their questioning techniques in class, some teachers reported that many of the questions that they used were closed, requiring only one-word or "yes or no" answers, and expressed a desire to change this tendency (Black et al., 2003). AfL teachers view questioning as a valuable tool to promote thinking and rich classroom discourse, which provides a wealth of information by which to evaluate their students' understanding of the subject matter (Black et al., 2003). Therefore, training to recognize and utilize questions that produce lengthy student responses is crucial. Question types are also an important area of CLIL research; studies have shown that CLIL teachers use more open than closed questions when inquiring about facts (Pascual Peña, 2010). This finding is important, as one of the main aims of CLIL is to establish meaningful classroom discourse to improve language competence and develop oral communication skills (Maljers et al., 2001). Therefore, CLIL teachers and AfL

teachers share the common goal of using questions to promote expression and effective communication.

2.6.2.c Wait time

Another important element in effective questioning is providing students with sufficient time after the question has been asked to form an answer (Black et al., 2003). Leaving wait time allows students to use critical thinking skills to formulate a response without assuming that one will be provided to them by the teacher. The Kings, Medway, Oxfordshire Formative Assessment Project (KMOFAP) determined that providing a wait time of less than one second discourages students from participating in classroom discourse, and the average wait time that most teachers used was calculated to be 0.9s. Conversely, increasing wait time promotes positive effects, including: increasing the length of answers; decrease in failure to respond; and students challenging or improving answers of other students (Rowe, 1974).

An extension of this strategy involves giving students the opportunity to discuss their answers in pairs or groups for 30 seconds before responding. The teacher may also "wait and recap," meaning wait until students have mentioned all key words and then recap their answers in a short summary. This approach supports the teacher's role as the mediator in the classroom rather than the authority, and also prompts students to take more responsibility for their own thinking and learning.

2.6.2.d Interpreting student response

Student answers are a valuable resource for teachers and provide evidence sometimes used to shape curriculum design. In addition to diagnostic questioning, other question types may be used to measure student progress, gaps in knowledge and motivation as the unit progresses (Black and Wiliam, 1998a). Questioning also prompts students to have a more active role in the classroom as participants rather than being passive entities (McCallum, Hargreaves & Gipps, 2000). Once an answer has been given, the teacher can diagnose the students' grasp of the concept. Incorrect answers should be viewed as an opportunity to begin class discussions rather than being dismissed. This leads to an atmosphere where being wrong is not a cause of embarrassment, and encourages students to offer input despite not being

completely certain of their responses. Feedback from a teacher or peer impacts student self-confidence, which is discussed further in the following section.

2.6.3 Feedback

Feedback is one of the most important elements in assessment for learning, as it focuses on developing understanding and treats mistakes as learning opportunities (Clark, 2011). In other words, feedback acts as a way of showing students frequent errors or misunderstandings and teaching them how to change their approach so as to correct these errors and maximize learning (Cauley & McMillan, 2010).

The use of effective feedback is a key factor in helping students cross their Zone of Proximal Development (ZPD). Vygotsky's theory of the Zone of Proximal Development describes a threshold that students have allowing them to work up to a certain ability level. Working alone, the student is able to complete activities only up to this level. However, with intervention, or assistance from a teacher or mediator, the student's ZPD is crossed, allowing them to work at a higher ability level (Vygotsky, 1978). This makes it possible for students to work at a higher level during group collaboration or while receiving help on a task from their teacher through effective feedback.

The intervention occurs as the teacher gives oral or written feedback to students. The traditional approach to providing feedback involves grading and giving comments on summative assessments. However, this approach provides little more than error correction. Even if the student did choose to revise and learn based on error corrections, the assessment window has been closed. The same rule applies for writing. Rather than focusing on the students' errors, the focus should be on what they did well and what can be done to improve.

Tunstall and Gipps (1996) developed a typology for teacher feedback including two major categories: the first relates to socialization and the second relates to assessment. Socialization feedback refers to the teacher giving students messages regarding respectful behavior toward their peers in the classroom. Assessment feedback is a multifaceted and contains four major types: A, B, C and D. Table 2.3

indicates the assessment types, and I have noted those specific to AfL in the "AfL Focus" column.

Table 2.3 Types of Feedback

Type	**Feedback Given**	**AfL Focus**
A1	Rewarding	√
A2	Punishing	
B1	Approving	√
B2	Disapproving	
C1	Specifying Attainment	
C2	Specifying improvement	√
D1	Constructing achievement	
D2	Constructing the way forward	√

(Adapted from Tunstall & Gipps, 1996: 394)

Interpreting the third column, feedback in AfL classrooms relies on rewarding students for good work and encouraging improvement rather than punishment for poor work or focusing on mistakes. Effective feedback places the teacher in the position of a reflective agent providing students with meaningful guidance to enable attainment of learning goals (Pollard, 1990). To accomplish this, teachers offer feedback not only on an exam at the end of a lesson, but as a continuous form of intervention to enable students to improve their performance on a daily basis.

Some feedback may be directed toward the class as a whole. However, if possible, individual feedback should be offered in the form of meetings each month or trimester with each student to inform them of their progress in relation to the learning objectives (Sutton, 1995). While keeping track of each student may be daunting, there are simple ways to monitor students that are less time consuming. A teacher may select one or two students in each class to observe more carefully during the lesson. At the end of the lesson, the teacher writes down observations and uses this in addition to a select portfolio of work to give feedback to both the

student and the parents (Sutton, 1995). Parents or caretakers have been cited as integral to the learning process (Townsend, 1997) and should also be involved in the feedback loop.

The development of assessment practice by each teacher plays a role in promoting motivation and achievement in students, and feedback is crucial in this process. Several guidelines for good assessment and feedback practice are listed below.

Good assessment and feedback practice should:

1. Help to clarify what good performance is (goals, criteria, standards)
2. Encourage 'time and effort' on challenging learning tasks
3. Deliver high-quality feedback information that helps learners to self-correct
4. Provide opportunities to act on feedback (to close the gap between current and desired performance)
5. Ensure that summative assessment has a positive impact on learning
6. Encourage interaction and dialogue around learning (peer and teacher-student)
7. Facilitate the development of self-assessment and reflection in learning
8. Give choice in the topic, method, criteria, weighting or timing of assessments
9. Involve students in decision-making about assessment policy and practice
10. Support the development of learning groups and communities
11. Encourage positive motivational beliefs and self-esteem
12. Provide information to teachers

(Spiller 2009: 15-17)

Effective feedback emphasizes what students have done well, which especially benefits lower achievers since they are less likely to receive positive comments and more likely to receive error correction (Young, 2005). To avoid this pattern, teachers employ comment-based feedback that includes highlighting the positive aspects of the work rather than mistakes. In addition to telling the students what

they have done well, AfL teachers also provide information about how to improve their performance. Continuous feedback through intervention during the lesson keeps students on track to achieve goals and provides scaffolding (structured comments meant to help students arrive at the answer that they are searching for) to close the learning gap (Young 2005).

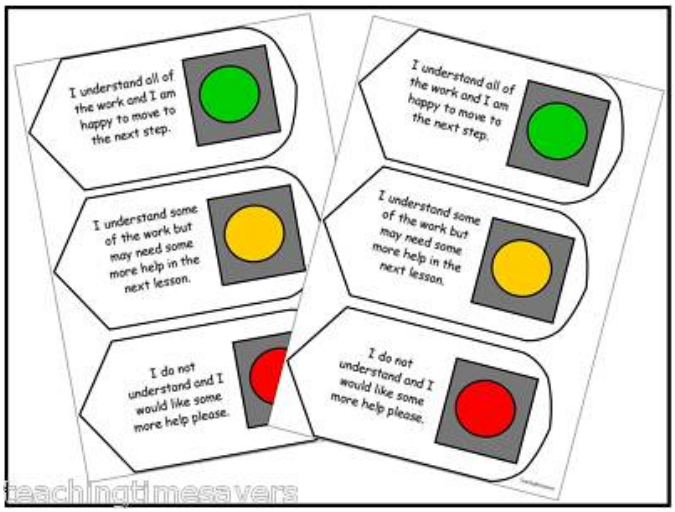
2.6.4 Self-assessment

A key, though frequently overlooked, learning tool that plays a role in returning ownership of learning to learners is self-assessment. According to Leitch et al. (2007) students seem to lack an active voice in the classroom, inhibiting participation. Self-assessment is a way to return the voice to students while developing metacognitive abilities and self-esteem (Bingham, Holbrook and Meyers, 2010). Challenges arise with self-assessment because it is not intuitive; students must be trained before being able to evaluate their own work (Bingham, Holbrook and Meyers, 2010). To train students on how to self-assess, some teachers provide additional materials such as rubrics or checklists and exemplars (examples of stellar work) to which students may compare their own work to effectively self-assess (To and Carless, 2015).

Self-assessment builds autonomy by training students to recognize what constitutes "good work" to attain the best results possible for their own work. This is one of the key tenants of AfL theory- making inferences and planning for further steps (Black and Wiliam, 1998a, 1998b). To build autonomy, students must have intrinsic, or internal, learning goals that they work toward attaining. Therefore, the task of the teacher becomes providing a risk free learning environment that promotes the emotional and psychological health of the students (Clark, 2011). Achievement of this goal involves training students to successfully self-assess, which for younger learners often requires the integration of mediating artifacts, such as posters or signs. These methods were put into practice by the AfL teachers involved in the present study, and are therefore described in detail.

Some methods of self-assessment techniques in the AfL classroom are traffic lights, thumbs up/ down and smiley faces. All three tools are used in a similar way: after the teacher has presented the content of the class, they ask students whether they have understood the basic concepts or if more explanation is necessary. In the case of traffic lights, students hold up a green light, meaning that they have understood the concept and feel prepared to move on; a yellow light indicating that they have further doubts; or a red light, which indicates a need to dedicate more time to explanation of the material. Thumbs up/ down and smiley faces mimic this approach, substituting a sad face or thumbs down for a red light.

Figure 2.2: Traffic Lights [3]



All are simple and quick forms of self-assessment, which allow the teacher to instantaneously evaluate the students' comprehension of the material. In addition to engaging the students with the use of visual aids, traffic lights help students, especially lower achievers, to express their lack of comprehension in a way that does not expose them to judgment or embarrassment in front of their peers. In other words, it is far more desirable to hold up a red card than raising one's hand and saying, "I don't understand" in front of a group of peers. It creates a sense of

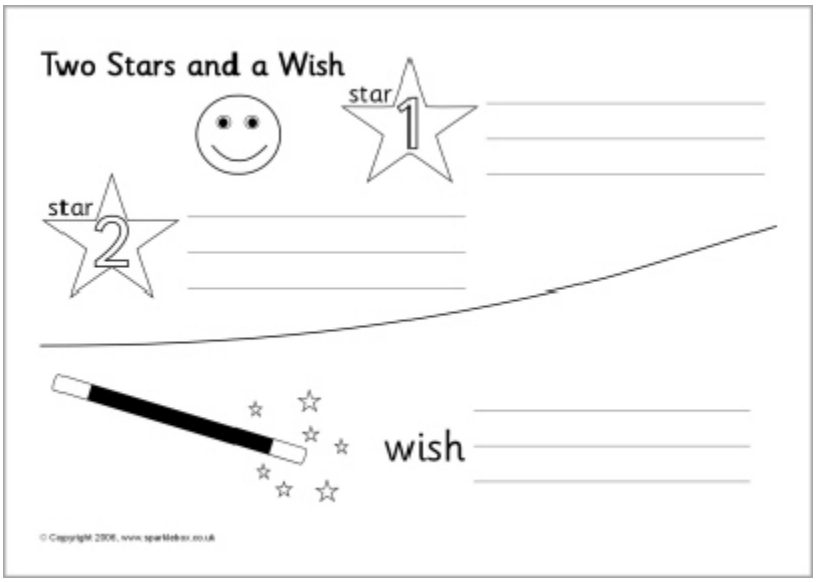
[3] Image taken from the-treasure-box.co.uk

equality or inclusion for students who feel on the margins of classroom learning (Leitch et al., 2007) and ensures the positive emotional and psychological circumstances are being maintained in the classroom environment (Clark, 2011).

While traffic lights, smiley faces and thumbs up/ thumbs down provide instantaneous self-assessment that take very little classroom time, these methods do not provide insight into the concepts that require further elaboration. In order to determine this reasoning, some teachers supplement traffic lights with meta-reflection at the end of each class or unit. This reflection involves asking students to write "minute papers" a brief summary of the topic at the end of the lesson allowing them to retroactively reflect on their knowledge (Angelo and Cross 1993). Keeping a regular reflection journal allows students to look back on what they have found difficult in the past, boosting their confidence as they see their progress in addressing each point.

AfL teachers also provide students with the opportunity for self-evaluation at the end of the lesson. One effective method involves providing a rubric of "I can" statements based on what they have learned throughout the unit or class. Students then put a tick next to they items they have mastered. If they cannot put a tick, it indicates both to the teacher and the student where the gaps in learning lie and where improvements must be made (Webb and Jones, 2009) At the end of the lesson, students complete "two stars and a wish" stating two positive things (stars) that they have done well during the class and an area for improvement (wish). The two stars and a wish technique enables students to evaluate their own work in an honest way and establish goals for future work.

Figure 2.3: Two Stars and a Wish [4]



Two stars: two positive aspects

A wish: area for improvement

Finally, reflection on a student's achievement is encouraged by the creation of a portfolio (Sutton, 2005) of their own self-selected exemplar work. The portfolio is accompanied by a short explanation as to why they have selected certain assignments as well as a reflection on the progress that has been made (Angelo and Cross, 1993). Once students are able to track this progress and see the results of their effort, it makes daily activities more meaningful.

2.6.5 Peer assessment

Many teachers underestimate students' capacity to act as a mediator to their peers. However, working in groups to provide feedback and assessment can benefit both higher and lower achieving students. For the higher achievers, the ability to provide feedback and help to their peers engages them in the learning process. It has been shown that those who give help to their fellow students generally benefit the most from the activity (Askew and Wiliam, 1995). For the middle and lower achieving students, peer correction sessions provide them the ability to see "exemplar work" or examples of high caliber assignments completed by their peers. Exposure to

[4] Image taken from sparklebox.co.uk

exemplar work allows lower achievers to get an idea of where their work needs to be to obtain better results. Looking at examples of peer work, students recognize what constitutes "good work" (Sadler, 1989; To and Carless, 2015) and, with training and help from the teacher, are able to create criteria and standards based on their observations (Black and Wiliam et al., 2008). Additionally, peer assessment encourages students to take a more active role in the classroom (Flórez & Sammons, 2013) and has been proven to improve metacognitive skills. A study by Hodgson and Pyle (2010) noted that integrating traffic lights in order for student to assess their peers' work was beneficial in this regard.

Peer assessment sessions may provoke criticism, especially by parents that believe the teacher is taking shortcuts in the interest of saving correction time. To avoid this stereotype, the teacher should encourage students to provide peer feedback rather than simple error correction. To facilitate the process, AfL teachers provide students with assessment criteria in advance, encouraging reflection and focusing on areas of improvement rather than errors. Once students have the opportunity to engage in peer assessment regularly, it becomes second nature and they develop the ability to provide scaffolding to their fellow students. As they continue to engage in discussions comparing their work to others, students enhance their knowledge of what "good work" involves (To and Carless, 2015), improving the quality of their output.

2.6.6 The formative use of summative tests

As previously mentioned, summative assessment has been criticized (Black et al., 2003; Stiggins, 2007; Shohamy, 2001) as an insufficient means of measuring students' abilities or knowledge, and potentially detrimental to learners' motivation. However, as assessment researchers Harlen and Crick stated in a 2002 report reviewing the impact of summative testing on student motivation, the combination of summative and formative assessment has the potential of improving standards.

> If the claims for both summative assessment and formative assessment are valid, the two could co-exist in educational practice, combining to raise

standards for all students (Harlen and Crick, 2002:10)

Using the two assessment practices effectively involves giving post-test feedback, sometimes through peer and self-assessment, to revisit challenging questions or concepts. AfL teachers take the opportunity to review test results as a way of detecting areas of improvement and focusing their teaching on these areas (Black et al., 2003), perhaps giving students the opportunity to re-take the test once the gaps have been closed.

2.7 Criticism of AfL

While AfL has been praised as having a positive impact by researchers and teachers alike, the approach has still endured criticism. Some concerns include time constraints and the amount of resources and planning a teacher must dedicate to making AfL work. Implementing assessment for learning involves a complete change in teaching philosophy and classroom culture (Black et al., 2003), therefore there is a need for best practice in developing uniform strategies that can be used and shared by teachers (Aberger, 2010). To make such changes, teachers need extra planning time and support. Increased wait times after questions, peer and self-correction sessions and increased verbal feedback also leads to a slower paced class and concerns regarding adequate coverage of content (Kirten et al., 2007). As a consequence, concerns have been raised regarding national testing standards being met properly.

Another problem for teachers in implementing AfL is the conflicting agenda of summative and formative assessment, which has the potential of creating tension between teachers, administrators and parents. One of the greatest challenges lies in the fact that formative assessment is informal, not systematic (Poehner and Lantolf, 2005), and therefore requires a shift in beliefs as to what assessment is meant to accomplish. This is especially true in countries such as China, where testing is a top priority and may determine a student's entire life course. There have been cases of complaints against teachers who have attempted to implement AfL by parents and administrators who believe that this approach interferes with curriculum test

standards. These critics do not recognize the broader aims of AfL, which are to engage students in "real learning" and not just cultivating the ability to achieve on tests (Reay and Wiliam, 1999).

In addition to changing the classroom culture, teachers also have to adapt their pedagogical approach, which can be a stressful and uncertain process. In contrast to many classrooms where the teacher is viewed as an authority, AfL requires a more student-centered approach, which many teachers find to be risky (Kirton et al, 2007) because it forces them to "let go" (Black et al., 2003) and allow students to take responsibility to promote active learning. The sustainability of AfL has also been called into question, as many teachers report the need to be aware of its implementation at all times (Kirton et al., 2007). The level of commitment to AfL has a tendency to wane over time, especially as new groups of students arrive each year and teachers must start from the beginning and introduce them to AfL fundamentals (Kirton et al., 2007). Without continuous support from fellow teachers and administrators, as well as access to additional funding for resources and time allowances, AfL is difficult to sustain, leading some teachers to abandon the approach altogether. Nevertheless, it has proven to have a positive impact on student achievement and motivation. This has prompted a movement toward its implementation as a means of enhancing the overall assessment process to benefit the learner. Research findings on the effects of AfL are presented in the next sections.

2.8 The evaluation of AfL by research

AfL continues to be significant in the field of education as evidenced by the ongoing research being conducted. In 1997, Paul Black and Dylan Wiliam conducted an extensive literature review to define formative assessment and determine its impact. After conducting a review of the literature, Black and Wiliam determined that some studies overestimated the effects of AfL based on small sample groups, (Black and Wiliam, 2003) but did find evidence of positive effects of the assessment technique. Since then, AfL practices have been integrated into the curriculum in

many countries, namely England and the United States. Subsequently, Flórez and Sammons (2013) conducted a second extensive search into AfL studies using a number of databases. For the purposes of this dissertation, a selection of AfL studies have been chosen based on the review by Black and Wiliam and the more recent literature review by Flórez and Sammons, as well as other subsequent representative studies. Sections 2.8.1-2.8.3 present literature on the effects of AfL on the following categories: achievement, motivation, and lower achieving students.

2.8.1 AfL and achievement

A number of empirical quantitative studies have concluded that the implementation of one or more AfL strategies leads to increases in student achievement (Kirton et al., 2007; Webb and Jones, 2009; Kellard et al., 2008). Many of these studies are based on the gains made by students on a pre-test to post-test basis, using an experimental group with exposure to AfL techniques and a control group. These techniques varied: some AfL students were engaged in daily self-assessment practice (Fontana and Fernandes, 1994); others were exposed to more frequent testing after which they were given feedback by an experienced teacher (Martinez and Martinez, 1992); and others involved adapting tasks to students' individual learning needs and closely monitoring progress (Bergan et al., 1991). In all cases, the experimental group showed significant improvement on the post-test compared to the control group, in some cases even doubling their scores.

In a longitudinal study, performed by a university professor who kept a record of his experience with 7000 students over the course of several years, AfL strategies involved the use of frequent summative testing punctuated by regular feedback. The students were required to demonstrate mastery by passing the exam, earning at least a 90% before moving on to the next task (Whiting, Van Burgh and Render, 1995). As a result, grade point averages increased as well as a complete change in the students' learning styles as a result of this incremental testing and feedback. As the year progressed, the time that it took students to master the material decreased, as did the number of retakes required to pass each exam. Qualitative studies of the

students' attitudes toward school in general also revealed positive feedback from students regarding formative assessment practices (Whiting, Van Burgh and Render, 1995).

While it could be argued that teaching style led to the positive changes, Whiting maintains that the AfL strategies could be used by any teacher to improve student achievement. In fact, research has shown that the implementation of these strategies proves to be more cost effective than class size reduction. Based on several studies in the United States, researchers concluded that the achievement difference in math and reading was higher in schools that implemented AfL than those which lowered the number of students per class (Yeh, 2009).

As a result of these findings, Black and Wiliam concluded that the use of formative assessment did, in fact, provoke significant gains in achievement, with an effect size of 0.4 to 0.7.

> Gains in achievement associated with formative assessment were among the largest ever reported for educational interventions. Specifically, students in classrooms with embedded formative assessment nearly doubled their rate of learning (effect size = 0.4 to 0.7), making 12 months of gains in 6 to 7 months (Black and Wiliam, 1998a: 36)

This seemingly small increase has the potential of increasing academic outcomes considerably. To put the numbers in perspective, if an effect size of 0.7 could be achieved nationally, it would have the potential of raising the international ranking in mathematics (based on attainment scores) of a country such as the United States into the ranks of top five countries such as Japan and Korea (Beaton et al., 1997). In short, the use of AfL in classroom practice has a demonstrated positive effect on student achievement.

2.8.2 AfL and student motivation

The shift that has occurred regarding classroom assessment has led researchers and teachers to focus on student emotions in relation to the assessment process. Part of

this focus has explored the effects of assessment and alternative assessment on student motivation. One such study conducted by Butler in 1988 sought to provide a link between teacher feedback and student motivation. Researchers divided 11 year-old students into three groups where they were asked to complete writing assignments. The feedback varied according to the group: students were either given a grade, a grade with written feedback or simply feedback on their work. They were then given the opportunity to improve their draft a second and third time, receiving the same type of feedback after each revision. Results were measured by comparing the grades students were given on each task. This indicated an increase in the feedback only group over the course of the three tasks, and a decline in the grade only and feedback with grade students. Researchers concluded that while feedback can often be helpful for students, it may be undermined by the negative motivational effects of giving grades, stating that preoccupation with results may lower the quality of student performance on tasks (Butler, 1988). Therefore, AfL seeks to de-emphasize the importance of grades, focusing instead on giving learners feedback to foster improvement.

A qualitative study measuring student motivation performed by Ross, Siegenthaler and Tronson (2006) yielded positive results. With a sample of 400 biology students from a university, teachers used several AfL techniques including continuous testing, feedback and self-assessment. Teachers gave students a series of tests at different intervals throughout the unit, asking the students to evaluate their own learning that had taken place in the interim. After being given several different tasks, including take-home assignments, and self-reflective creative written assignments, the students were encouraged to self-evaluate, and feedback was provided in a timely fashion by the teachers. In the end, 50% of the total grade was dedicated to formative assessment. To measure the students' response to these AfL techniques, researchers conducted focus groups and gave questionnaires regarding their beliefs about these techniques and how they were related to the biology content. The results indicated that students believed they benefited from the AfL techniques, which improved recall, understanding and higher cognitive outcomes,

as well as making the assignments more enjoyable (Ross, Siegenthaler and Tronson, 2006). The majority of students (95-98%) commented that the series of evaluations throughout the unit helped improve their recall and understanding of the material. Table 2.4 gives a small selection of the comments from the student focus groups. While the students involved in this study were of university age, their comments are representative of the positive effects of AfL in all levels of education.

Table 2.4 Effectiveness of assessment for learning

1. Students' self-evaluation showed that they benefitted from knowing the purpose and format of the assessment beforehand.

2. Students appreciated consistent, clear and unambiguous instructions about what they were supposed to do in the assessable work, what was expected of them and what they could expect as a result of their input.

3. Students liked the way assessments were spread throughout the semester. Representative comments indicated that this enabled them to keep the concepts they had learned 'fresh in their minds for longer' and prompted them 'to study on a more regular basis'.

4. Students felt that the feedback they received from their assessments was excellent; a typical comment was "feedback really helped me learn the topic in question".

5. This was a common theme; students from the 2004 cohort were amazed they could still remember fine details 12 months later. They contrasted this with other topics/subjects where they had attempted rote learning and now could not remember details.

6. Interactive learning strategies prompted students to use self-assessment and study approaches that were different from those they habitually used in the past, e.g. using games and computer quizzes more often.

7. Students reported that the examination results they got were better than expected, or as expected. No students in the focus groups reported being disappointed with their final result.

(Ross, Stiegenthalter and Tronson, 2006)

The results indicate that through interactive strategies such as self-assessment exercises and feedback, students were able to connect more with the material they were studying. Continuous evaluations meant that they studied more regularly

rather than waiting until the last minute for a final exam. As a result, the students were able to recall details of the material at a later date. Setting learning goals and getting regular feedback ensured that the students were satisfied with the end result of their final evaluation, motivating them to move on confidently knowing the material and repeat the process (Ross, Siegenthaler and Tronson, 2006). The recent introduction of technology into the learning environments, specifically the tablet, has also given teachers the means to offer students immediate feedback, which students have reported as beneficial (Dekkers et al. 2014). This study illustrates the effectiveness of feedback on learners irrespective of age and the content studied.

The final example, selected for comment due to the fact that the students were from primary school and of a similar age to those in the present study, was conducted with 44 students aged 9 or 10 from the United States. Students were divided into four groups: in two of the groups, graduate student instructors stressed learning goals (learning how to solve problems), while in the other two the instructors set performance based goals (simply solving problems). At the end of the lessons, one group from each goal set evaluated their problem-solving capacities while the other group completed an attitudinal questionnaire regarding their work. The results, which were based on skill, motivation and self-efficacy, indicated that students who were given performance goals without self-evaluation demonstrated the lowest motivational scores of all possible groups. These results show that the effects of self-evaluation outweighed the overall goal type, though the study also suggests that stressing learning goals leads to increased achievement and motivation (Schunk, 1996). By setting continuous learning goals and working toward improvement, teachers are emphasizing the correlation between achievement and effort rather than achievement and biology (Aberger, 2010). The effects of this distinction mean that AfL can empower students to take control over their learning, thus building confidence and sustaining motivation (Yin et al., 2008).

2.8.3 AfL and lower achieving students

According to Black and Wiliam (1998b) students expected to benefit the most from AfL strategies are lower achievers. For these students, poor results on exams lead them to believe that success is impossible, regardless of the output of effort. As mentioned previously, placing a large amount of pressure on grade attainment has the potential of lowering the quality of performance (Butler, 1988). Other research argues that many students are unfairly labeled as having learning difficulties when really this might not be the case; the use of AfL has been shown to minimize this prejudice. In fact, the effects of AfL on achievement have proven to be the greatest with lower achieving students, English language learners and special needs students (Boston, 2002). These gains were primarily the result of creating a classroom culture of learning that includes continuous feedback and allows students to self-assess their understanding and address areas of improvement, encouraging progress toward learning goals (Yin et al., 2008).

In an investigation by Bergan et al. (1991), researchers divided 838 children from disadvantaged areas of the United States and during an eight-week course to track their progress while using AfL strategies. Instructors were trained to implement a system starting with a diagnostic examination to determine the students' level followed by continuous assessment throughout the course punctuated by student-teacher consultation. The students were presented with a variety of tasks enabling teachers to assess their skills and progress, with instructors catering to the needs of each individual student.

In the control group, on average 1 child in 3.7 was referred as having particular learning needs and 1 in 5 was placed in special education, the corresponding figures for the experimental group were 1 in 17 and 1 in 71. Based on these high numbers, researchers concluded that the potential of many children in conventional teaching settings is underestimated. Some may even be ridiculed unnecessarily, having their futures prejudiced and perpetuating this cycle of failure (Bergan et al., 1991).

Rather than judging students on their performance on tests, in an AfL setting the goal becomes closing the gap between higher and lower achievers and raising attainment overall (Black and Wiliam 1998a). The approach has been shown to offer a new pedagogy, actively involving students in the assessment process, which improves motivation and self-esteem (Black and Wiliam, 1998a). Research also suggests that students who understand what they are being asked to learn and provided with a model for success are more likely to set learning goals (Young 2005). Above all, the introduction of these strategies creates a sense of equality and inclusion, namely for students who normally feel on the margins of classroom learning, such as lower achievers (Leitch et al., 2007).

2.9 Chapter summary

Chapter 2 began with an overview of the shifting views of assessment that have occurred over the past several decades, followed by a general description of the methods of assessment found in the classroom. A distinction was then made between summative assessment and assessment for learning. The effects of testing on students, which drove the need for alternative ways to assess and the emergence of AfL, were then explored. This was followed by a description of AfL techniques identified by researchers and the ways in which teachers employ such techniques to empower learners' improvement through assessment. The chapter concluded by reviewing relevant literature on AfL and its effects on student achievement, motivation and lower achieving learners. The next chapter focuses on the role of motivation in L2 learning.

Chapter 3: Theoretical Framework: Motivation in second language learning

3.1 Introduction

While motivation is difficult to define, there is nothing more important in sustaining a long-term language-learning goal (Guilloteaux and Dörnyei, 2008). This chapter gives a working definition of motivation based on the work of several leading figures in L2 motivation research (Williams and Burden, 1997, Gardner, 2010) and examines different factors that influence and enhance motivation when learning a foreign language. It then examines the educational shift that took place in motivational research in the 1990's and the development of a more practical application of techniques for measuring motivation in the classroom. The chapter concludes by discussing motivational strategies used by teachers in the L2 classroom as well as the development of the Motivated Orientation of Language Teaching (MOLT) scheme.

3.2 Defining motivation

According to researchers and teachers, motivation is one of the most crucial factors in learning a second language over an extended period of time (Guilloteaux and Dörnyei, 2008). The difficulty of defining motivation lies in that it is not a tangible or measurable quality and visible only through its effect on learners (Dörnyei, 2001). While most teachers and students would agree that motivation is a key element that can lead to success or failure in language learning (Dörnyei and Csizer, 1998), many might find it difficult to articulate what exactly this entails.

Robert Gardner, pioneer of research into L2 motivation, states that giving a simple definition of motivation is impossible: it is only possible to list characteristics of a motivated student. This student is goal directed, persistent, attentive, has desires (wants), exhibits positive affect, has expectancies, demonstrates self-confidence (self-efficacy), and has reasons (motives) for learning (Gardner, 2007). The study of what motivates learners contains many conflicting theories, but the one thing that researchers agree upon is that motivation is one of the key factors in initiating

interest in language learning and sustaining this interest for a prolonged period of time.

> Without sufficient motivation, even individuals with the most remarkable abilities cannot accomplish long-term goals, and neither are appropriate curricula and good teaching enough to ensure student achievement. (Dörnyei and Csizér, 1998: 203)

According to Williams and Burden, researchers responsible for developing a framework for motivation in language learning in the 1990's, motivation has become a general term to describe a disposition toward learning (1997). In reality, it is a highly complex concept influenced by several different factors (Gardner, 2010). The Williams and Burden framework is concerned with internal factors leading to motivation (interest in the activity, perceived value in activity, self-concept, and many more) as well as external factors (significant others such as parents, teachers and peers, the learning environment and the broader context) (1997). Extensive investigation has identified factors needed for an individual to become motivated and sustain this drive. Williams and Burden (1997) offer a general definition based on their findings:

Motivation:

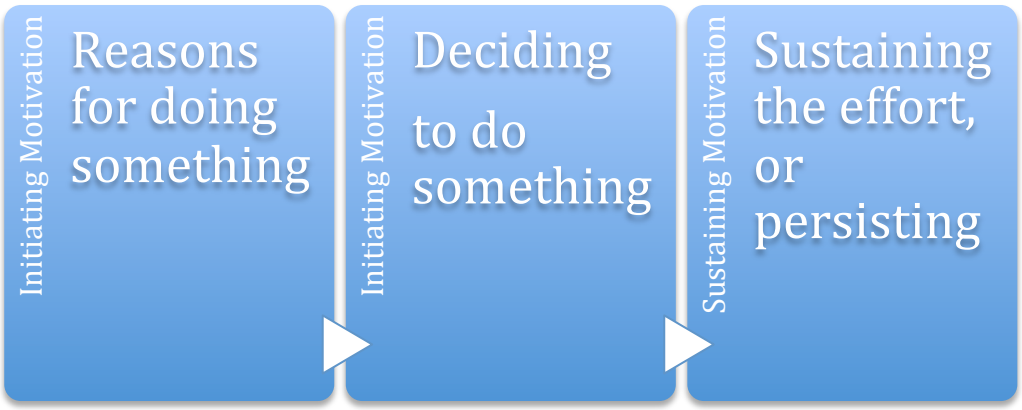
- A state of cognitive and emotional arousal
- Which leads to a conscious decision to act, and
- Which gives rise to a period of sustained intellectual and/or physical effort
- In order to attain a previously set goal (or goals)

(Williams and Burden, 1997: 120)

The model for motivation contains three parts. First, an individual must have reasons for doing something. Next, they must make a conscious decision to do this activity. Finally, the individual must sustain the task or activity for a prolonged

amount of time. While the first two stages are concerned with *initiating motivation* the last stage encompasses *sustaining motivation*.

Figure 3.1: A three-stage model of motivation



(Williams and Burden, 1997: 121)

The factors leading to initiating motivation are numerous and are discussed at length in the next sections. While early models attribute motivation to an instinct, need, or drive (that the learner might not necessarily be aware or in control of), this line of thinking shifted with cognitive views of motivation, which stipulate that beliefs or information residing in the mind of the learner are what lead to the conscious drive to act (Ames and Ames, 1984). These mental processes could be related to information-seeking, metacognitive and cognitive strategies, emotional states or self-evaluations (Ames and Ames, 1984).

The next sections examine several models of motivation that have been developed by researchers and psychologists over the past several decades. Due to the high volume of research in this area, the chapter focuses on a timeline of key models that have had a significant impact on the field and hold relevance to the present study. This includes those focusing specifically on the study of L2 motivation in the field of education, which contributed to the development of the MOLT observational scheme (Guilloteaux and Dörnyei, 2008), a framework applied in the methodology

of the present dissertation.

3.3 Models of second language motivation

While psychologists, linguists and educators have been questioning the origins of motivation in the field of psychology for years, concrete theories related to second language learning motivation began to develop in 1950's. Since that time, dozens of theories have emerged and been debated regarding what motivates students to learn a second language. It should also noted that, despite prolific research, scholars have yet to come to one conclusion on the source of L2 motivation, as stated by Masgoret and Gardner:

> Many times, these hypotheses have been proposed based on results obtained by different researchers, in very different contexts, and often using different measures. Because of all the differences between these studies, it is difficult to arrive at unequivocal answers concerning their validity. (Masgoret and Gardner, 2003)

Masgoret and Gardner's statement begs the question of how the field of L2 motivational research can connect differences in context, measures and participants to draw conclusions to benefit the learner. When extended into the educational context, the challenge addressed is how teachers may employ strategies to initiate and sustain motivation as the difficult process of learning a second language progresses. This section summarizes important work in the development of the field and justifies the model chosen for the study.

The field of L2 motivation research has gone through three major stages, starting with Robert Gardner's work in the 1960's, in which Gardner argued that the most important drive toward learning a second language is the desire to integrate with others in the L2 community (Gardner and Lambert, 1972). Gardner's work, rooted in social psychology, was called into question when cognitive motivational theories became popular over the next several decades (Dörnyei, 2003). Models to be discussed include: achievement motivation theory, attribution theory, goal setting/

goal orientation theories and self-determination theory as well as Dörnyei's work regarding the L2 Motivational Self System, which considers visualizing an ideal future self as integral to maintaining L2 motivation. In the 1990's, researchers across the world began to apply these theories to L2 classroom learning in an effort to understand what motivates students in these settings and develop concrete teaching strategies and task based motivational tools (Dörnyei, 1998).

3.3.1 Gardner's socio-educational model (Integrativeness)

Gardner and Lambert's research cites language as a mediating device between two different cultural communities (1972). Funded by the Canadian government, the research sought to provide insight into the complex co-existence of English and French speakers. The theory is based on the concept of *integrativeness*, which refers to an "openness to identify, at least in part, with another language community" (Masgoret and Gardner, 2003), which Gardner identifies as the most crucial aspect driving a learner to master a foreign language. In other words, learners who feel the need to integrate into another culture have more of a drive to learn a language than learners who do not.

To apply this theory to empirical study, Gardner created the Attitude/ Motivation Test Battery (AMTB) (1985), an instrument that assesses integrativeness based on three categories: attitudinal variables, motivational variables toward learning a second language and integrative orientation. Attitudinal variables are concerned with the learner's attitude toward the learning situation, with a distinction between attitude toward the teacher and the course (Gardner, 1985). Motivation is divided into three subcategories: motivational intensity; desire to learn the target language; and attitudes toward learning the target language (Gardner, 1985). The integrative orientation assesses the individual's desire to meet, socialize with and become friends with members of the target language community. Additionally, in his work, Gardner includes sub-tests to measure integrative and instrumental orientations, which reflect the distinction between learning a language to interact with members of another community (integrativeness) versus learning for the benefit of furthering

one's career goals or for other practical reasons (instrumental) (Masgoret and Gardner, 2003). Variables for Gardner's AMTB are shown in table 3.1.

Table 3.1 Gardner's AMTB (1985)

Variable	**Sub-Categories**
Attitudes Toward the Learning Situation	Evaluation of the course
	Evaluation of the Teacher
Integrativeness	Attitudes toward the target language group
	Interest in foreign languages
	Integrative orientation
Motivation	Motivational intensity
	Attitude toward learning the target language
	Desire to learn the target language
Orientations	Instrumental orientation
	Integrative orientation

(Gardner, 1985: 9)

The attitude section of the AMTB is administered in the form of a questionnaire in which participants are given a series of statements that may be positive or negative (e.g. "I would like to learn more about French Canadians" "I hate French") and asked to give their opinion on each item based on a seven-point scale (strongly disagree, moderately disagree, slightly disagree, neutral, slightly agree, moderately agree, strongly agree).

The motivational component of the AMTB is a separate section based on multiple-choice questions with three possible answers that are graded on a scale of 1-3 (e.g. Compared to my other courses, I like French a. the most (3) b. the same as others (2) and c. least of all (1)). Based on students' responses to the questionnaire, Gardner

was able to generate empirical evidence of their level of motivation to learn the second language.

The AMTB has since been used in studies in numerous cultural contexts. However, some have pointed out the difficulty of adapting the questionnaire to other cultural environments where, unlike in Canada, there is little exposure to the target language outside of the classroom setting (Dörnyei, 1990; Cid, Graneña and Tragant, 2009). In addition, while Gardner's theories remain some of the most important in modern motivational research, they have come under scrutiny by others who cite that integrativeness, while now a mainstream term, has no clear roots in motivational psychology and may be difficult to define (Dörnyei, 2003). Gardner himself has admitted that the term may mean different things to different individuals (Gardner, 2001). Thus, researchers began to suggest alternative theories that might affect learner motivation.

3.3.2 Self-determination theory: intrinsic versus extrinsic motivation

Arguably one of the most influential models of motivation is self-determination theory (Deci and Ryan, 1992). A substantial amount of research has gone into understanding the distinction between extrinsic motivation, which is based on earning rewards or avoiding punishment, and intrinsic motivation, or the individual's desire to perform the task for its own sake (Bénabou and Tirole 2003). The theory is based on the reasons that *move* an individual toward wanting to do something (Ryan and Deci, 2000), which could be externally or internally based. In the classroom environment, extrinsic motivation comes from grades, teacher approval, parents and most notably, passing an exam (Williams and Burden, 1997). Intrinsic motivation, on the other hand, develops from factors such as an interest or curiosity or an innate need for autonomy (Ryan and Deci, 2000). Some researchers consider extrinsic and intrinsic motivation mutually exclusive (Harter, 1981) while others believe that the two are almost always linked (Williams and Burden, 1997).

The question that often arises is: which type is more powerful? Many business environments seem to believe that extrinsic motivation has a greater impact on

employees, and therefore offer payment incentives and benefits to push workers toward greater output. However, studies have shown that while these rewards might entice employees at the beginning, they do not help individuals sustain long-term motivation (Bénabou and Tirole 2003) and the most crucial factor is innate interest in the task. In a learning environment, intrinsic motivation has been shown to result in higher quality learning and creativity (Ryan and Deci, 2000). Expressing interest or curiosity in the subject matter can be enough to sustain long-term learning without the need for reward. In a controversial piece of research by Deci, Koestner and Ryan (2001) it was suggested that tangible rewards may affect intrinsic motivation negatively. In the classroom setting, students who are given more autonomy by parents and teachers report increased intrinsic motivation, curiosity and desire for a challenge (Deci, Nezlek and Sheinman, 1981). This theory also applies to professional contexts, where research has shown that employees may be more motivated by completing meaningful projects with impact than achieving by financial gain (Pink, 2009; Bénabou and Tirole, 2003).

3.3.3 Achievement motivation

While Gardner argues that integrativeness is one of the greatest factors for individuals to learn an L2, the need for achievement may be equally important (Atkinson, 1964). The achievement motivation model is based on conflicting factors that drive an individual to approach or avoid a certain task or situation. Positive factors include whether the learner *expects* to be successful in performing the task, the *value* that the learner places on the task and the *need for achievement*. The negative factors include: *fear of failure*, the *incentive to avoid failure* and the *probability of failure* (Raynon, Atkinson and Brown, 1974). Achievement motivation is determined by the relative strength of these opposing forces (need for success, fear of failure) against one another as the individual is compelled to approach or avoid a given task (Williams and Burden, 1997). For example, the positive or negative reaction that a student might have toward high-stakes testing. While the student may feel highly motivated to succeed in order to achieve a higher educational status or appease parents and teachers, the stress of the impending

examination might provoke an avoidance reaction. In the end, overall motivation is determined by the strength of these two individual factors (Williams and Burden, 1997).

3.3.4 Goal theories

The approach/ avoidance concept related to achievement theory has undergone several changes, leading researchers to explore the importance of goals as a catalyst for motivation (Elliot and Dweck, 1988). According to Covington, goals give learners purpose and guide behavior:

> All actions are given meaning, direction and purpose by the goals that individuals seek out…quality and intensity of behavior will change as these goals change (Covington, 2000: 174)

Establishing goals has always been an important part of the learning process, though the concept of a goal might be more complex than it appears (Williams and Burden, 1997). Researchers and cognitive psychologists have made distinctions between different types of goals, which may be used for different purposes. The most well known distinction is between goals that are *performance versus mastery* goals (Ames, 1992) or *performance versus learning* goals (Dweck and Leggett, 1988). Performance goals refer to a desire to not seem intellectually inferior or avoid embarrassment in front of peers. Learning or mastery goals are based on an authentic desire to learn the material (Williams and Burden, 1997). Furthermore, learning goals are generally associated with a higher level of intrinsic motivation, whereas performance goals suggest a lower sense of intrinsic motivation (Heyman and Dweck, 1992). The importance of establishing goals should not be underestimated and has been noted as a key factor in assessment for learning practice (Black and Wiliam, 1998a).

3.3.5 Attribution theory

Attribution theory was the first to successfully challenge Atkinson's achievement theory, becoming a popular model throughout the 1980's. The theory, developed by

Bernard Weiner, hypothesizes that learners are scientists or researchers and constantly analyze the results of past events (Weiner, 1992). This analysis, in turn, has the potential to affect future events, as learners are inclined to use results obtained in order to determine how to proceed. Weiner claims that an individual's past successes or failures is followed by the motivation to initiate a future action (Weiner, 2000). For instance, a student might take a standardized test to determine which university they will ultimately attend. If the student scores poorly on the exam, they may decide to forego attending university. Conversely, a student who scores well on the exam might decide to apply to a more prestigious university or rigorous program (Weiner, 2000). There are four main factors to which learners are likely to attribute failure or success: ability, effort, task difficulty and luck (Weiner, 1992).

In this case, the success or failure may be attributed to the exam, though when repeated failure occurs it is more likely for a student to attribute the failure to themselves (Weiner, 2000). As noted in Chapter 2, assessment reformer Rick Stiggins points out that high stakes testing has the potential of placing students on a "winning or losing streak" and the more times a student fails, the more likely they are to become de-motivated, affecting future successes (Stiggins, 2007), leading to a dangerous cycle. Dörnyei argues that attribution theory is especially relevant to L2 motivation, due to the high frequency of language learning failure worldwide (Dörnyei 2003). A study conducted by Williams, Burden and Al-Baharna regarding attribution theory in L2 motivation produced thought-provoking findings. These included a marked difference in factors to which teachers and students attributed successes and failures (in relation to classroom performance or achievement of a certain grade) in the language-learning context. The findings also indicated that students attributed their success in language learning to internal factors (e.g. practice) and their failures to external factors (e.g. inadequate teaching methods) (Williams, Burden and Al-Baharna, 2001).

3.3.6 The L2 Motivational Self System

The importance of the self-concept of the learner in creating a sense of agency (Williams and Burden, 1997) has long been acknowledged as one of the keys to understanding motivation. However, the question of the inner drives responsible for motivating learners still remains unknown. While Gardner maintained integrativeness is the most powerful catalyst, research by Dörnyei and Csizer suggested that Gardner's conceptualization was limited, after conducting a large-scale motivational study in Hungary (2002). This claim led to further research into the traditional psychological concept of the *self* and how it relates to motivational drives to learn an L2. Thus, a new theory known as the *L2 Motivational Self System* (Dörnyei, 2005) was developed.

Research on the *self* as related to L2 motivation is rooted in work that began in the 1980's. The notion of multiple "selves" was first introduced by Markus and Nurius, who distinguished between three possible selves: the *ideal self*, or what the learner wants to become; the self that the learner could *possibly* become; the self that the learner is *afraid* of becoming (1986). Shortly after, Higgins expanded upon this idea, theorizing that there are two types of future selves that might drive the present self toward L2 proficiency. The first is the *ideal self*, which represents the attributes that the learner would like to possess (Higgins, 1987), and if proficiency in a target language is included in these attributes, a learner may have more desire to learn a second language (Dörnyei, 2009). The second, the *ought to self*, refers to what the learner feels are the attributes that they should or ought to possess (Higgins, 1987). Unlike the ideal self, the *ought to self* represents duties or obligations imposed by external sources (Macintyre, Mackinnon and Clément, 2009). With self-determination theory in mind, the *ideal self* is driven by intrinsic motivation and the *ought to* self by extrinsic motivation. Dörnyei draws upon the *ideal* and *ought to selves* to formulate the L2 Motivational Self System, adding a third factor: *L2 Learning Experience*, which represents the impact of the students' learning environment (Dörnyei, 2005).

Table 3.2: The L2 Motivational Self System

The L2 Motivational Self System
1). *Ideal L2 self:* The L2 specific facet of one's 'ideal self': if the person we would like to become speaks an L2, the *ideal L2* self is a powerful motivator to learn the L2 because of the desire to reduce the discrepancy between our actual and ideal selves. Traditional integrative and internalized instrumental motives would typically belong to this component. 2). *Ought-to L2 self* concerns the attributes that one believes one *ought to* possess to meet expectations and *avoid* possible negative outcomes. This dimension corresponds to Higgins's *ought self* and thus to more extrinsic (i.e. less internalized) types of instrumental motives 3). *L2 Learning experience* concerns situated, 'executive' motives related to the immediate learning environment and experience (e.g. the impact of the teacher, the curriculum, the peer group, the experience of success). (Dörnyei, 2005: 29)

Following Dörnyei's development of the L2 Motivational Self System, several studies were conducted, starting with that of Dörnyei and Csizer in Hungary (2002) and expanding to encompass four different countries and nationalities (Japanese, Iranian, Korean, Chinese). A compilation of these results can be found in *Motivation, Language Identity and the L2 Self* (Dörnyei, 2009). Nevertheless, research into the L2 Motivational Self System is still at the beginning stages and further studies must be completed in order to understand possible selves as a framework for L2 motivation (MacIntyre, MacKinnon and Clément, 2009).

3.4 A shift in L2 motivational research: toward an education based model

The previous sections have discussed models of motivation, explored the effects of motivation in developing an *ideal self* and various factors that motivate students to learn a second language. The next sections provide a basis for empirical study on L2 motivation, which has been ongoing since the late 1950's. Throughout the past several decades, there have been various approaches to measuring motivation, both

quantitative and qualitative, though until the 1990's almost all approaches were informed by Gardner's model and sought to link motivation and achievement (Ushioda, 2001). The research question that Gardner originally set out to answer was:

> How is it that some people can learn a second language so easily and do so well while others, given what seem to be the same opportunities to learn, find it impossible? (Gardner and Lambert, 1972)

Researchers are still trying to find an answer to this question, though over the years the methods have become more sophisticated. For many years, the primary concern was creating working definitions of motivation and theories related to factors that influence learners. However, the creation of these theories was not being tested or used to promote motivation in students (Guilloteaux and Dörnyei, 2008). Furthermore, Gardner's social psychological approach, which was by far the most popular in motivation research, did not directly address the classroom implications of motivating learners and did not provide teachers with explicit tools to achieve this goal. In the 1990's, an "educational shift" occurred when researchers started looking at the classroom environment itself as having a strong influence on motivation, as indicated by Dörnyei:

> Part of the revival of interest in L2 motivation in the 1990s was prompted by a large number of studies that attempted to reopen the research agenda with a 'new wave' educational focus (e.g. Brown, 1990, 1994; Clement, Dornyei & Noels, 1994; Crookes & Schmidt, 1991; Dörnyei, 1994; Julkunen, 1989, 1993; Oxford & Shearin, 1994; Skehan, 1989, 1991; Ushioda, 1994, 1996; Williams, 1994). This new movement cannot be specifically tied to any particular school or scholars because, as the above list shows, a number of researchers in different parts of the world appeared to come up with similar ideas at around the same time. (Dörnyei, 1998: 124)

In an effort to make a distinction between language learning motivation and classroom learning motivation (Gardner, 2007), researchers began to develop tools

and means of conducting empirical studies that addressed classroom learning. Since this shift occurred, new methodological developments have emerged based on theoretical advances and, as a result, motivational research began to achieve a new level of maturity (Guilloteaux and Dörnyei, 2008).

The forces at work in the classroom are multilayered, and include components specific to the course, the teacher and the learner group (Dörnyei, 2003). Dörnyei set out to find strategies for teachers to apply to motivate their students to learn a second language in a classroom setting based on extensive research. The following sections focus on the research undertaken to identify these strategies, the strategies themselves and how they may be observed in the classroom context.

3.5 Motivational strategies used by teachers in the language classroom

Dörnyei sought to explore the drives that motivate learners to pursue second language learning, drawing from well-known motivational theories. The result was a three-level framework containing: *language level* (integrative and instrumental subsystems), *learner level* (need for achievement, self confidence), and *learner situation level* (consisting of course-specific, teacher specific and group specific motivational components) (Dörnyei, 1994).

Table 3.3: Components of Foreign Language Learning Motivation

Language Level	Integrative Motivational Subsystem Instrumental Motivational Subsystem
Learner Level	Need for Achievement Self-Confidence -Language use anxiety -Perceived L2 self-competence -Causal Attributions -Self-Efficacy

Learning Situation Level **+Course Specific Motivational Components**	Interest Relevance Expectancy Satisfaction
+ Teacher Specific Motivational Components	Affiliative Drive Authority Type Direct Socialization of Motivation -Modelling -Task Presentation -Feedback
+Goal Specific Motivational Components	Goal-Orientedness Norm & Reward System Group Cohesion Classroom Goal Structure

(Dörnyei, 1994: 280)

Based on this framework, Dörnyei created a list of 30 macro strategies designed to help teachers better motivate their students throughout the lesson. Each of these strategies was then divided into micro strategies, which in the end offered over 100 strategies teachers could apply to motivate their students (Dörnyei and Csizér, 1998). Noting the overwhelming task for teachers of filtering through such an extensive list, Dörnyei created a more manageable list of ten recommendations entitled: "Ten Commandments for Motivating Language Learners". However, the synthesis of rules, commandments and recommendations by Dörnyei came after scrutiny for a lack of empirical evidence in order to support such claims (Gardner and Tremblay, 1994). Therefore, to ensure the validity of the list, a study was

conducted by Dörnyei and Csizér (1998) to create a revised Ten Commandment list based on strong empirical backing.

Participants included 200 Hungarian teachers of diverse locations and educational settings who were asked to complete two questionnaires. The first aim was for teachers to rate each motivational strategy on a seven-point scale based on perceived importance, and the second to rate the same strategies based on frequency of use in the classroom (Dörnyei and Csizér, 1998). Each participating teacher was given only one of the two questionnaires to complete, with data analysis consisting of compiling a ranking system of all 51 strategies included in the questionnaire. The top ten strategies on the list were then considered the definitive "Ten Commandments".

Table 3.4: Ten Commandments for motivating language learners

1. Set a personal example with your own behavior.
2. Create a pleasant, relaxed atmosphere in the classroom.
3. Present the tasks properly.
4. Develop a good relationship with the learners.
5. Increase the learners' linguistic self-confidence.
6. Make the language classes interesting.
7. Promote learner autonomy.
8. Personalize the learning process.
9. Increase the learners' goal-orientedness.
10. Familiarize learners with the target language culture.

(Dörnyei and Csizér, 1998: 215)

Results from the study and the initial list of 100 strategies compiled by Dörnyei were then published as the full book *Motivational Strategies in the Language Classroom* (2001). This publication was intended to be a tool for teachers and contained more than 100 concrete motivational strategies divided into three phases: creating basic motivational conditions; generating initial motivation; maintaining and protecting motivation; and encouraging positive self-evaluation (Dörnyei, 2001).

After years of documenting teacher's self- reports of what constitutes effective motivational strategies and their use in the classroom, Dörnyei identified the need for more objective empirical evidence documenting motivational practice used by these teachers (Guilloteaux and Dörnyei, 2008). Thus, the Motivated Orientation of Language Teaching (MOLT) classroom observation scheme was created; a tool used by researchers to measure observable L2 motivational strategies used during the lesson. The scheme, which is the instrument used to document motivational strategies in AfL and non-AfL classrooms in this dissertation, is discussed in the following section, as well as its practical application to empirical research.

3.6 The MOLT classroom observation scheme

The MOLT scheme is a hybrid of Dörnyei's (2001) motivational strategies framework for foreign language classrooms and the Communication Orientation of Language Teaching (COLT) framework, which uses the principle of real time coding to measure each variable (Spada and Frölich, 1995). The framework was developed as a means of measuring classroom motivational strategies being used by the teacher as the lesson progresses. These strategies are based on instructional techniques implemented by the teacher as well as self-regulating strategies used by students to increase their own motivation (Guilloteaux and Dörnyei, 2008). After identifying the key classroom motivation strategies, the scheme was created to measure how these strategies were implemented in real time.

The five categories that the framework encompasses are: *learners' motivated behavior; teacher discourse; participation structure; activity design and encouraging positive retrospective self-evaluation.* The first category, learners' motivated behavior, is related to the level of engagement, attention and participation that learners demonstrate throughout the lesson, while the other four categories, based on Dörnyei's model of motivational teaching practice (2001), are related to the strategies the teacher uses during the lesson to motivate students. Each category is divided into sub-categories (see table 3.5) along with a description (Guilloteaux and Dörnyei, 2008).

Table 3.5 MOLT classroom observation scheme

Attention	Learners' Motivated Behavior	Students appear to be paying attention: They are not displaying any inattentive or disruptive behavior; they are looking at the teacher and following his or her movements, looking at visual stimuli, turning to watch another student who is contributing to the task, following the text being read, or making appropriate nonverbal responses
Participation	Learners' Motivated Behavior	Students are actively taking part in classroom interaction or working on assigned activity.
Volunteering for a teacher-fronted activity	Learners' Motivated Behavior	At least one third of the students are volunteering without the teacher having to coax them in any way.
Social Chat	Teacher discourse	Having an informal (often humorous) chat with the students on matters unrelated to the lesson.
Signposting	Teacher discourse	Stating the lesson objectives explicitly or giving retrospective summaries of progress already made toward realizing the objectives.
Stating the communicative purpose or utility of an activity	Teacher discourse	While presenting an activity, mentioning its communicative purpose, its usefulness outside the classroom, its cross-curricular utility, or the way it fits into the sequence of activities planned for the lesson.
Establishing relevance	Teacher discourse	Connecting what has to be learned to the students' everyday lives (e.g., giving grammatical examples with references to pop stars).
Promoting integrative values	Teacher discourse	Promoting contact with L2 speakers and cultural products and encouraging students to explore the L2 culture and community
Promoting instrumental values	Teacher discourse	Highlighting the role that the L2 plays in the world and how knowing the L2 can be potentially useful for the students themselves as well as their community.
Arousing curiosity or attention	Teacher discourse	During the presentation of an activity, raising the students' expectations that the upcoming activity is going to be interesting and/or important (e.g., by asking them to guess what they are going to do next, or by pointing out fun, challenging, or important aspects of the activity or contents to be learned).

Scaffolding	Teacher discourse	Providing appropriate strategies and/or models to help students complete an activity successfully (e.g., the teacher thinks aloud while demonstrating, reminds students of previously learned knowledge or skills that will help them complete the activity, or has the class brainstorm a list of strategies to carry out the activity).
Promoting cooperation	Teacher discourse	Setting up a cooperative learning activity, or explicitly encouraging students to help one another, offering suggestions on how best to do this.
Referential questions	Teacher discourse	Asking the class questions to which the teacher does not already know the answer, including questions about the students' lives.
Promoting autonomy	Teacher discourse	Offering students a choice of activities, involving them in making decisions regarding the timing of an activity, having them use the Internet or do research (e.g., for oral presentations, projects, and displays).
Group work	Participation structure	The students are mingling, working in fluid pairs, or working in groups (simultaneously or presenting to the whole class).
Pair work	Participation structure	The students are working in fixed pairs (simultaneously or presenting to the whole class).
Tangible reward	Activity design	Offering students tangible rewards (e.g., candy, stickers) for successfully taking part in an activity.
Element of interest, creativity or fantasy	Activity design	The activity contains ambiguous, paradoxical, problematic, controversial, contradictory, incongruous, or exotic material; connects with students' interests, values, creativity, fantasy, or arouses their curiosity (e.g., predict-and-confirm activity).
Personalization	Activity design	Creating opportunities for students to express personal meanings (e.g., experiences, feelings, opinions).
Intellectual challenge	Activity design	The activity presents an intellectual challenge (e.g., it involves a memory challenge, problem or puzzle solving, discovering something, overcoming obstacles, avoiding traps, or finding hidden information).

Tangible task product	Activity design	The students are working on the production of a tangible outcome (e.g., a poster, a brochure).
Individual competition	Activity design	The activity involves an element of individual competition.
Team competition	Activity design	The activity involves an element of team competition.
Neutral feedback	Encouraging positive retrospective self-evaluation	Going over the answers of an exercise with the class without communicating any expression of irritation or personal criticism.
Process feedback	Encouraging positive retrospective self-evaluation	Focusing on what can be learned from the mistakes that have been made, and from the process of producing the correct answer.
Elicitation of self or peer-correction	Encouraging positive retrospective self-evaluation	Encouraging students to correct their own mistakes, revise their own work, or review/correct their peers' work.
Effective praise	Encouraging positive retrospective self-evaluation	Offering praise for effort or achievement that is sincere, specific (i.e., more than merely saying "Good job!"), and commensurate with the student's achievement. N.B.: Ability feedback ("You are very good at English") or praise involving social comparison ("You did better than anyone else in the class") is not recorded as effective praise.
Class applause	Encouraging positive retrospective self-evaluation	Celebrating a student's or group's success, risk-taking, or effort by applauding (either spontaneously or following the teacher's lead).

(Guilloteaux and Dörnyei, 2008: 62-64)

The following section discusses the MOLT scheme and how it was applied to empirical research in a 2008 study by Guilloteaux and Dörnyei. The MOLT scheme was chosen as an instrument in this dissertation because it was the first framework to empirically measure teacher's motivational practice based on classroom observation rather than self- reporting questionnaires.

3.7 Practical applications of the MOLT

The MOLT scheme was based on a culmination of years of research into the motivational strategies used by teachers in L2 classrooms. It was constructed out of a desire to observe how these strategies were used and measure their impact on student motivation. The study was conducted in 20 junior high state schools in South Korea. Participants included 27 language teachers and a total of 1,381 students in 40 different classes.

The research questions of the study:

1. How does the teacher's motivational teaching practice affect the students' classroom motivation in terms of the level of their attention, participation, and volunteering?

2. What is the relationship between the students' self-reported motivation (assessed by questionnaire), their actual classroom behavior, and the teacher's classroom practice?

Three instruments were used to identify the strategies used by teachers and how they affected student motivation: *(a) a classroom observation scheme (b) a student questionnaire and (c) a post-lesson teacher evaluation scale*. The classroom observation scheme used was the MOLT scheme. The questionnaire was based on the Student Motivational State Questionnaire [Appendix 6] intended to measure general attitude and motivational factors. There were 20 items on the questionnaire to be rated on a six-point scale (1= definitely not, 6= totally true). The post lesson teacher evaluation scale was intended to provide a post-hoc evaluation of the teacher's disposition. It focused on motivational features of the teacher's behavior and a 1-6 point ranking system (e.g. radiates enthusiasm (6) unenthusiastic (1)). (Guilloteaux and Dörnyei, 2008).

After an extensive testing phase comprised of a pilot study of eight EFL classes (with a total of four teachers and 293 students), adjustments were made to the classroom observation scheme and questionnaire. In the main study, the procedure was conducted as follows: students completed the motivational questionnaire before the first period of the morning or afternoon. Then, the researcher observed the lesson

using the MOLT framework to make note of which techniques were used on a minute-by-minute checklist, calculating the duration of time the teacher sustained each one to determine overall classroom motivational teaching practice. At the end of the lesson, the researcher discussed the results with the teacher in order to ensure that both were in agreement with the use of strategies, making adjustments when necessary. The post-hoc teacher evaluation was completed by the researcher at the end of the lesson (Guilloteaux and Dörnyei, 2008).

To analyze the data collected using the MOLT scheme, researchers added up the sum of the minutes (ranging from 0-45) during which a specific student behavior, motivational technique or activity took place and entered them into an SPSS data file. A Cronbach alpha reliability coefficient was used to measure internal consistency among the motivational features. Results indicated a highly significant positive correlation between teacher's motivational practice and students' motivated behavior during the lessons. The study also found that teacher's motivated practice affected 40% of the variance in students' self-reported motivated behavior (i.e. the questionnaires), which researchers noted to be a remarkable result, considering all of the other factors that could affect student behavior in class (Guilloteaux and Dörnyei, 2008).

While this framework is the main instrument used to measure teacher's motivational behavior in this dissertation, some adjustments were made to both the instrument itself and the method, which are discussed in greater detail in Chapter 5 (Methodology).

3.8 Chapter summary

The chapter began by providing a working definition and a brief introduction on motivation. Several models that have been proposed by researchers were then presented, which were integral to the development of Dörnyei's strategies for teachers in motivating students to learn an L2. The shift in motivational research toward a more practical model related to education was then explained, as well as the desire to identify and classify strategies used by teachers to enhance student

motivation. The chapter went on to describe the MOLT scheme, a classroom observational tool that allowed researchers to observe these strategies as they were happening in real time. Finally, the practical applications of the scheme were described by giving an overview of Guilloteaux and Dörnyei's 2008 study. Chapter 4 continues by giving an overview of Appraisal theory and describing the framework created by Martin and White, which is used as an instrument in this study.

Chapter 4: Theoretical framework: Student Metacognitive Linguistic Strategies: The Use of Appraisal

4.1 Introduction

This chapter is concerned with student's use of linguistic strategies to convey stances, opinions or beliefs, known in Systemic Functional Linguistics (SFL) as APPRAISAL. The system of APPRAISAL originated in the interpersonal metafunction in Halliday's theory of SFL. Researchers Martin and White extended the system by creating a framework for APPRAISAL from the early 1990's, detailed extensively in their work *The Language of Evaluation in English: Appraisal in English* (2005). Martin and White's theory is rooted in exploring interpersonal language, namely the subjective stance of a writer or speaker as they present their personal stance in a text or discussion (2005). This stance can be made toward material (i.e. a work of art, an oral presentation, a piece of literature) or a person, and may be positive, negative or neutral. According to Llinares, Morton and Whittaker, the interpersonal function of language is essential to language assessment.

> SFL studies on the role of language at different educational levels show that knowing how to use the language for interpersonal purposes also plays an important role in school success (e.g. Christie, 2002, Coffin 1997; Schleppegrell, 2004). According to these studies, many students, both native and non-native speakers of the target language, need support in their use of interpersonal resources (for example, modality). For this reason, the development of the interpersonal function constitutes one of the key elements in the evaluation of English as a second language (Polias, 2003: 221) (Llinares, Morton and Whittaker, 2012: 221)

This chapter discusses the origins of Appraisal theory and its place within the context of SFL. Martin and White's framework for APPRAISAL is then described, focusing on the categories of JUDGEMENT and APPRECIATION. This is followed by a discussion of how these two categories specifically relate to AfL and benefit students in acquiring the resources necessary to voice their opinions on a specific

person, performance or piece of work. The chapter concludes by considering how APPRAISAL is used by students when evaluating their own performance, classroom environment and experience with the target language.

4.2 Origins of Appraisal Theory

APPRAISAL is a form of evaluation given by an author or speaker regarding attitude, the strength of these feelings and the ways in which these opinions or values are sourced (Martin and Rose, 2003). Development of Appraisal theory originated from work on narrative genres at the University of Sydney in the 1980's regarding the use of interpersonal meaning (the ways in which individuals enact social relationships) in the development of genres. Years later, a framework was developed in response to a growing need to extend the interpersonal model already in place, in conjunction with the Write it Right literacy program (Martin and White, 2005). Part of the Disadvantaged Schools Project, this writing program focused on considering lexical resources used to judge behavior in secondary school writing (Iedema, Feez, and White 1994). While the origins of Appraisal Theory focused on measuring the use of APPRAISAL features in students' writing, the framework has since been applied to measuring the ways in which children evaluate emotional responses to certain situations and experiences through spoken discourse (Mills et al., 2014). The present study aims to measure how students may also apply these strategies to self-assessment and evaluating their learning experience.

In past research, APPRAISAL has been applied to writing (Derewianka, 2009; Coffin, 1997; Iedema, Feez and White, 1994) and public and academic discourse (Fairclough, 1992, Lemke, 1992, Martin, 1995). Research on APPRAISAL has been conducted in classroom discourse (Llinares, Morton and Whittaker, 2012) and teacher talk (Chu, 2014), though application of the theory to primary classroom discourse has been limited. Traditionally, the ways in which APPRAISAL has been applied in the classroom setting is by tracking students' abilities to create an identity in their writing (O'Donnell, 2013, McCabe and Whittaker forthcoming) and

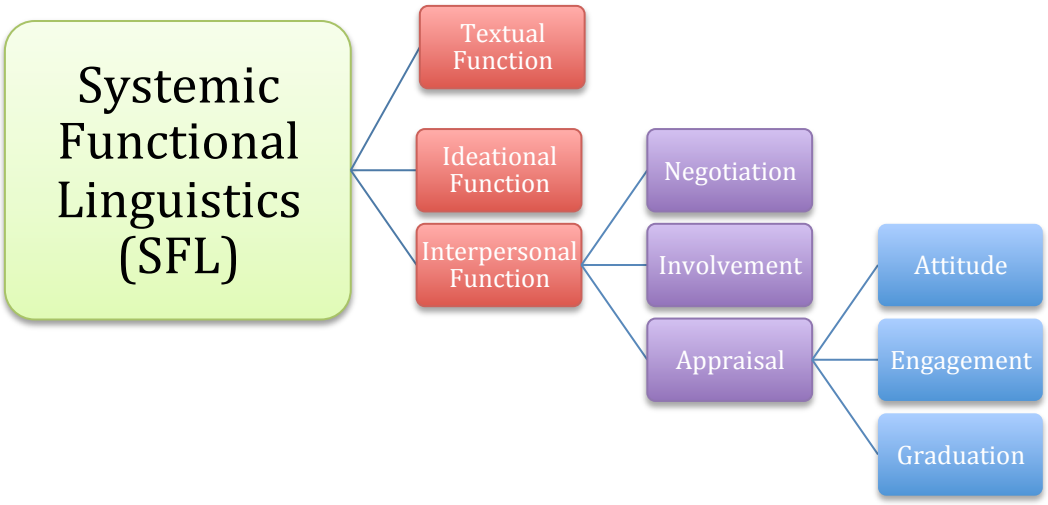
develop a voice used to persuade the reader of a certain point of view (Coffin, 2002, Derewianka, 2007).

When considered as a means of metacognitive analysis of students' beliefs on classroom practice and evaluating their own abilities or performance, APPRAISAL is a potentially powerful tool for both teachers and researchers. Therefore, the aim of this part of the study is to analyze certain APPRAISAL features, namely JUDGMENT and APPRECIATION, and how they are applied by students to self-assess their learning and performance, as well as evaluating their teacher and classroom experience.

4.3 Appraisal Theory in the context of SFL

APPRAISAL originated as part of M.A.K. Halliday's theory of systemic functional linguistics (SFL), which seeks to develop an understanding of how language and learning are interrelated (Halliday, 1993). This theoretical model focuses on the ways speakers create meaning through language, mainly, though not limited to, social and educational contexts. Figure 4.1 shows the placement of APPRAISAL within the structure of SFL, under the interpersonal function.

Figure 4.1: The placement of appraisal in systemic functional linguistics



Halliday, 1994

SFL identifies three distinct modes of meaning that operate at the same time in any situation in which language is spoken: the *ideational, interpersonal* and *textual.* The

ideational refers to how we make sense of our own experience (what's going on, who's doing what to whom); the interpersonal refers to the way in which we use language to construct relationships, interact and express feelings; and the textual acts as a link between the two by constructing sequences of discourse that flow and have cohesion and continuity (Halliday and Matthiessen, 2004). These general kinds of meaning are referred to as metafunctions (Halliday, 1994), encompassing what language conveys and how.

APPRAISAL is just one of three linguistic areas that make up interpersonal meaning, along with involvement and negotiation (Halliday, 1994). APPRAISAL is concerned with the ability of a speaker or writer to provide a subjective presence; to adopt a stance or make a claim that seeks to criticize, praise, approve or disapprove of a particular person or event (Martin and White, 2005). The categories of Appraisal are described in the following sections.

4.4 Categories of APPRAISAL

Martin and White recognize three APPRAISAL categories: ATTITUDE, ENGAGEMENT and GRADUATION with each containing several sub-categories (2005). ATTITUDE encompasses different options for expressing positive and negative evaluations, or in other words, different ways of feeling (e.g. Today was a very *sad* day for me, I am very *unhappy*). These feelings are related to three main systems--emotions, ethics or aesthetics—expressed by the speaker or writer. ENGAGEMENT, on the other hand, is concerned with the ability to engage with other voices and alternative positions. In some instances, this is related to reported speech (e.g. Chomsky *claimed to have shown* that…) though the author may also use ENGAGEMENT to give an opinion (e.g. I *contend* the fact of the matter is…). Finally, the third category, GRADUATION, provides linguistic resources that allow the speaker to soften (e.g. I'm *kind of* upset by what you said) or intensify (e.g. He's a *true* friend) an attitude or feeling (e.g. somewhat, slightly, rather, very, entirely, sort of, kind of) (Martin and White, 2005).

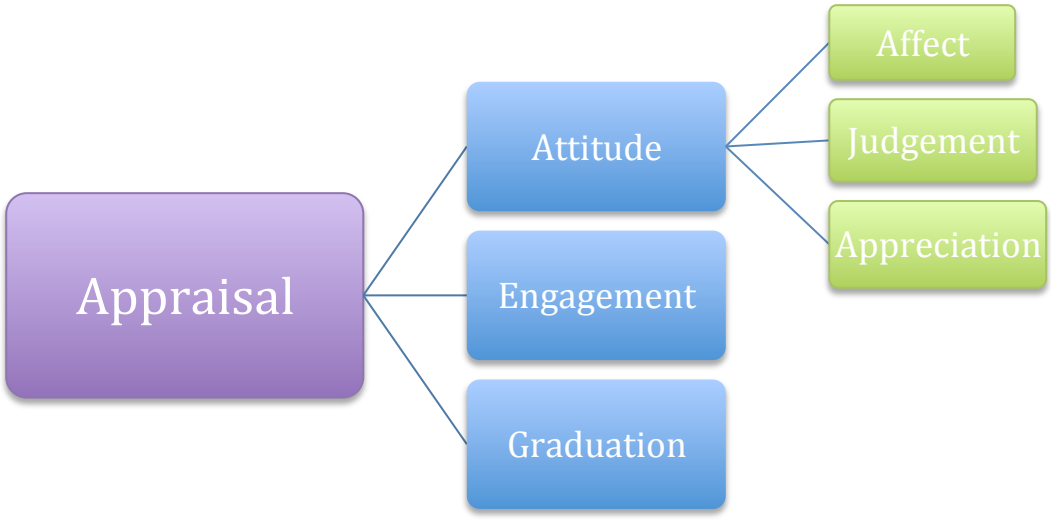
These resources are used to empower the speaker or author with the ability to declare a position that is strongly aligned or less strongly aligned with the values of

the community at large (Martin and White, 2005). The framework for APPRAISAL is highly complex with each category containing numerous sub categories. Therefore, the descriptions are meant to encompass the general meaning and not a comprehensive view. For a detailed, in-depth overview of all categories, refer to Martin and White (2005) or Chapter 2: Appraisal of Martin and Rose (2003). This study is concerned with the category of ATTITUDE, as it contains the linguistic elements necessary for students to assess people (themselves, their teacher or their classmates) as well as things or ideas (performances, presentations, tasks).

4.4.1 Attitude

As mentioned previously, ATTITUDE is a system related to mapping out kinds of feelings expressed by a writer or speaker and contains three sub-categories: AFFECT, JUDGEMENT and APPRECIATION. Traditionally, these areas are referred to as emotion (AFFECT), ethics (JUDGEMENT) and aesthetics (APPRECIATION) based on the type of feelings that they are meant to convey (Martin and White, 2005).

Figure 4.2: Placement of ATTITUDE under APPRAISAL

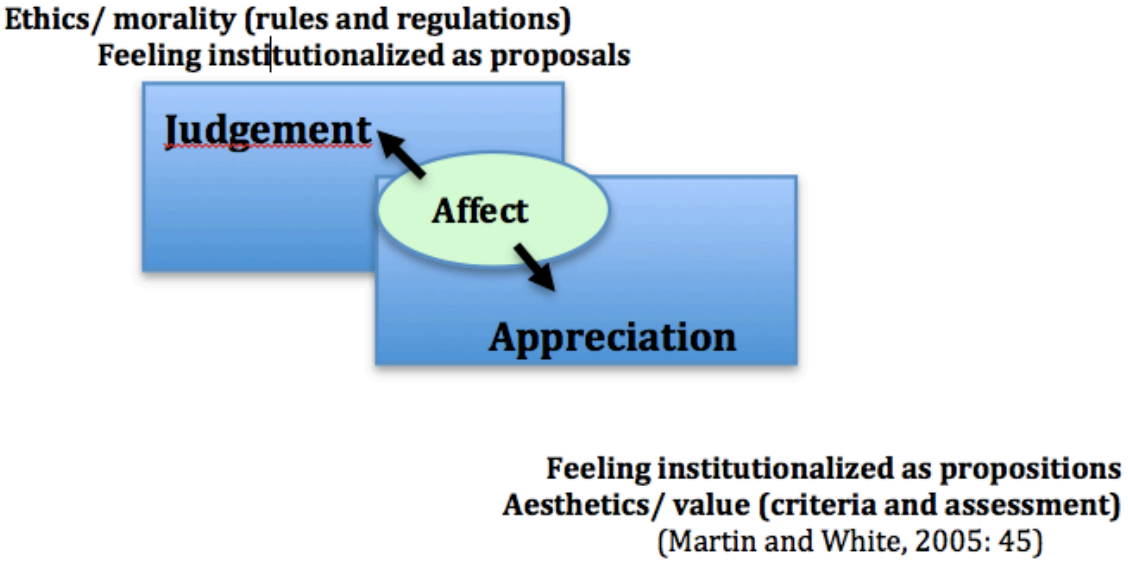


(Martin and White, 2005: 38)

While emotions are the focus of all linguistic resources related to ATTITUDE, AFFECT is concerned with expressing personal emotions, whether positive or negative. It seeks

to answer the question; "How are you feeling?" and can include a wide range of linguistic resources to register positive or negative feelings (e.g. "I'm so *bored* today," or "I was struck by a wave of *grief*") (Martin and White, 2005). While AFFECT is used to describe internal reactions to situations or circumstances, JUDGEMENT encompasses attitudes toward a specific individual or group of people based on perceptions of their behavior or character ("George's actions were *callous* and showed a *lack of empathy* towards those around him") (Martin and White, 2005). APPRECIATION involves making an evaluation on the value of things or ideas. Martin and White show the interrelation of AFFECT, JUDGEMENT and APPRECIATION in an institutionalized way (see Figure 4.3).

Figure 4.3: JUDGEMENT and APPRECIATION as institutionalized AFFECT



To show the connection between JUDGEMENT and APPRECIATION, Martin and White describe them as: "institutionalized feelings which take us out of our everyday common sense world into the uncommon worlds of shared community values" (Martin and White, 2005: 45). In this world, JUDGEMENT describes feelings about behavior or people, whereas APPRECIATION reflects feelings towards the value of things (Martin and White, 2005).

This study is concerned with the ways in which students use JUDGEMENT and APPRECIATION as a means of self-assessment or when describing feelings toward their learning experience. While the concept of AFFECT is not discussed further, it should be noted that it does play an important role in connecting these two systems together.

4.4.2 Judgement

JUDGEMENT is concerned with expressing attitudes toward a certain individual or group of people. It is commonly used in everyday conversation when making an evaluation about an individual's behavior, and can be used as a linguistic tool in certain professions, such as law or psychology. When making a JUDGMENT, the speaker or writer can use positive or negative terminology depending on the feeling expressed. First, a distinction between types of JUDGEMENT, SOCIAL ESTEEM and SOCIAL SANCTION, must be made. SOCIAL ESTEEM is used commonly in daily life and may be expressed orally through gossip, chat, jokes and stories shared with intimate acquaintances, such as relatives or friends. On the other hand, SOCIAL SANCTION is mostly expressed in writing through laws, rules or decrees by the church and state, with the possibility of punishment if broken (Martin and White, 2005: 52).

4.4.2.a. JUDGEMENT: SOCIAL ESTEEM

Martin and White detail three sub-categories under SOCIAL ESTEEM: NORMALITY**,** which relates to how unusual a person is, CAPACITY, or how capable they are and TENACITY, or how resolute they may be (2005: 52). Table 4.1 lists examples for each category, showing how positive and negative JUDGEMENTS may be applied. These examples are illustrative and meant to be used as a basis to recognize others in this category.

Table 4.1: Examples of JUDGEMENT: SOCIAL ESTEEM

	Positive (admire)	**Negative (criticize)**
Normality "how special?"	Lucky Normal Cool Fashionable	Unlucky Odd Erratic Dated
Capacity "how capable"	Powerful Healthy Clever Educated Competent Successful	Weak Stupid Inexpert Uneducated Incompetent, Unsuccessful
Tenacity "how dependable?"	Brave Cautious Careful Reliable	Timid Impatient Weak Unreliable

Martin and White (2005: 53)

The examples in the table show how SOCIAL ESTEEM is used to evaluate an individual's behavior or character, and may be used in conversation.

4.4.2.b. JUDGEMENT: SOCIAL SANCTION

SOCIAL SANCTION, found mostly in writing, is divided into two categories: VERACITY and PROPRIETY. While veracity is related to how honest an individual may, propriety is concern with ethics (Martin and White, 2005: 52). Similar to SOCIAL ESTEEM, SOCIAL SANCTION, may also be expressed in a positive or negative way (praise or condemnation) based on the assessment of the individual's character or actions. Table 4.2 gives key words associated with SOCIAL SANCTION.

Table 4.2: Social sanction

Social Sanction	**Positive (praise)**	**Negative (condemn)**
Veracity (truth) "how honest?"	Honest Candid	Dishonest Deceptive
Propriety (ethics) "how far beyond reproach?"	Moral Law abiding Modest	Immoral Corrupt Cruel

(Martin and White 2005:53)

When making a distinction between SOCIAL ESTEEM and SOCIAL SANCTION, Martin and White offer an analogy: "too much negative ESTEEM and we may need to visit a therapist, too much negative JUDGEMENT and a lawyer may need to be called in" (2005: 53). This statement sums up the everyday use of SOCIAL ESTEEM and the more formal use of SOCIAL SANCTION.

4.4.3 Appreciation

The function of APPRECIATION is to place a value on a given performance or task. In the classroom setting, these performances can be student presentations or a students' response to a teacher's question or request. APPRECIATION may also be seen in the form of teacher feedback or praise. Teachers also may encourage students to provide feedback related to peer's performances or responses, namely while engaging in peer assessment. APPRECIATION is categorized into three types: REACTIONS to things (do they catch our attention; do they please us?), their COMPOSITION (balance and complexity) and their VALUE (how innovative, authentic, timely…) (Martin and White, 2005: 56).

Table 4.3 shows types of APPRECIATION and provides examples of each in both a positive or negative context.

Table 4.3: Types of appreciation

	Positive	**Negative**
Reaction impact ('did it grab me?')	Captivating Fascinating Intense Remarkable	Boring Dry Monotonous Unremarkable
Reaction: quality ('did I like it?')	Okay Beautiful Enchanting	Bad Ugly Revolting
Composition: Balance ('did it hang together?')	Balanced Symmetrical Consistent	Unbalanced Uneven Disorganized
Composition: Complexity ('was it hard to follow?')	Simple Clear Detailed	Extravagant Unclear Simplistic
Valuation ('was it worthwhile?')	Profound Innovative	Shallow Conventional

(Martin and White, 2005: 56)

When considering these sub-categories, Martin and White classify REACTION, COMPOSITION and VALUATION in terms of a three-step mental process of the way that we look at things. REACTION is related to affection (emotive- "it grabs me" and desiderative- "I want it"), COMPOSITION to perception (how we order things) and VALUATION to cognition (2005:57). Based on the type of evaluation, the speaker or writer draws from these three categories to explain how they are affected by a given thing or performance.

4.5 Making the Distinction Between Judgement and Appreciation

Analyzing APPRAISAL is a challenging process, as oftentimes a certain example can be interpreted in more than one way depending on the context. Coders may annotate a text differently, despite being experts in the field of APPRAISAL research (Read, Hope and Carroll, 2007). When analyzing spoken discourse, which is spontaneous and sometimes ambiguous or incomplete, especially in the case of young learners, analysis is oftentimes challenging. Therefore, in the case of JUDGEMENT and APPRECIATION, Martin and White distinguish between these two categories, even though they may be used in conjunction. While JUDGEMENT represents the valuation of behavior, APPRECIATION represents the valuation of things. For example, a JUDGEMENT may be made of an artist and APPRECIATION is used to make a statement about the artist's work.

JUDGEMENT	APPRECIATION
a brilliant scholar	*a penetrating analysis*
he's a skillful player	*It was a skillful inning*

When considering JUDGEMENT and APPRECIATION in the context of textual or conversational analysis, it is essential to recall this distinction to correctly identify what the speaker is evaluating. In short, while JUDGEMENTS target people, APPRECIATIONS target things (Martin and White, 2005: 60).

4.6 Inscribed versus invoked

When it comes to applying this complex system, the stance may be expressed directly or indirectly, which Martin and White classify as INSCRIBED or INVOKED. The first type, INSCRIBED, contains a word that explicitly evaluates behavior (e.g. immoral, virtuous, just). The second type, INVOKED, represents a case in which there is no explicit word or phrase directly stating JUDGEMENT or APPRECIATION, but meaning can be derived indirectly (e.g. Bush delivered his inaugural speech as the United States President who collected 537,000 fewer votes than his opponent). In

this case, an indirect JUDGEMENT is being made on the president's capacity for public office based on the low number of votes that he received (2005).

Martin and White acknowledge that while an activity is appreciated as a thing, a JUDGEMENT of whoever accomplished it may be INVOKED, and the same in reverse (2005:67). For example, commentary on a football player's many errors implies an INVOKED JUDGEMENT of this player as a poor athlete. This process may involve creating a nominalization of a text displaying a JUDGEMENT by removing the agent, thus transforming the text into an APPRECIATION. Table 4.4 gives examples of the interactions between invocation and inscription in the category of ATTITUDE.

Table 4.4: Interactions between attitudinal invocations and attitudinal inscription

INSCRIBED JUDGEMENT and INVOKED APPRECIATION	**INSCRIBED APPRECIATION and INVOKED JUDGEMENT**
He proved a fascinating player	It was a fascinating innings (impact)
He played average (normality)	It was an average innings
He played strongly (capacity)	It was a strong innings

(Martin and White, 2005: 68)

Therefore, a hidden JUDGEMENT can be implied based on overt APPRECIATION and vice- versa.

4.7 Appraisal and AfL

APPRAISAL plays an important role in the classroom, and its use is encouraged in several AfL techniques. Considering the importance of administering feedback throughout the lesson and training students in peer and self-assessment techniques, it is possible that students exposed to AfL are more familiar with the use of APPRAISAL. When delivering feedback on a student's performance, the teacher uses APPRECIATION to encourage the student to recognize the quality of their work. This intervention from the teacher, if it is constructive (i.e. not punishing or disapproving without indicating a way to improve), encourages students to recognize what

constitutes good work and empower them with the ability to appraise their own performance or that of a peer. Therefore, one of the aims of the present study is to determine how AfL and non-AfL students convey stance when interviewed about their motivation in certain classroom situations.

4.8 Chapter Summary

Chapter 4 introduced APPRAISAL and its place in the SFL model. The chapter then outlined the three broad categories of this system, focusing on ATTITUDE, specifically JUDGEMENT and APPRECIATION. The two sub-categories of JUDGEMENT and APPRECIATION were described in terms of their relevance to society and social interactions under Martin and White's framework, which was used to code interviews of lower achieving AfL and non-AfL students. A distinction was made between the two sub-categories, as they are similar in nature and it is therefore important to determine what the speaker is evaluating to correctly label the APPRAISAL for analysis and interpretation. The chapter elaborated on the connection between the two, noting the ways in which JUDGEMENT can be invoked through APPRECIATION and vice-versa. Finally, the importance of APPRAISAL in the AfL classroom was discussed, and how it may be used by students when self-assessing, an important topic in this study. This is the final theoretical framework chapter--Chapter 5 continues by presenting the methodology of the study.

Chapter 5: Methodology

5.1 Introduction

Chapters 2-4 completed the review of relevant literature, establishing the framework for this study. Chapter 2 described methods of assessment, establishing AfL as a viable method for improving student achievement and detailing its role as a tool to reach learning goals. Chapter 3 gave a comprehensive overview of several approaches to motivational research, culminating in the identification of classroom L2 motivational techniques used by teachers. Finally, Chapter 4 provided a description of Appraisal Theory and its place in the Systemic Functional Linguistics model. The purpose of this chapter is to give the methodological approach, beginning with the research questions. The chapter continues by elaborating on the corpus, describing the participants and offering a description of the data and methods of collection. The chapter is divided into three parts following three analysis procedures that comprise this dissertation: Part 1: The use of L2 motivational strategies by teachers; Part 2: The effects of assessment on student self-reported motivation and feelings during the lesson; and Part 3: Lower achieving students' reflections on their own learning.

5.2 Methods and Research questions

The aim of this dissertation is to explore the relationship of AfL and L2 motivational strategies used by CLIL teachers, as well as their influence on learners' self-reported motivation and feelings in class. Furthermore, this study seeks to document the ways in which lower achieving students self-assess and comment on their learning experience. The study is divided into three parts with data sources and objectives for each. The data and analysis for the first part of the study focuses on teachers' use of motivational L2 strategies during classroom interactions; the second explores students' reported L2 motivation and feelings through questionnaires; and the third centers on the way lower achieving students comment on their learning in one-on-one interviews.

This is a mixed methods (MM) study, which takes into account quantitative and qualitative data. This mixed-methods approach provides a fuller picture and deeper understanding of the data (Johnson, Onwuegbuzie and Turner, 2007). MM research and triangulation, or using different sources to increase the validity of a study (Guion, Diehl and McDonald, 2002) is desired in the field of CLIL research (Pérez Cañado, 2012). Quantitative data are generated through measurement of the frequency and duration of L2 motivational strategies in the classroom and through the use of student questionnaires, while qualitative data are provided through the analysis of classroom and interview transcriptions.

The first part of the study uses classroom transcriptions to identify observable L2 motivational strategies. While studies have addressed the impact of certain AfL techniques on student motivation through the use of questionnaires and post-lesson interviews, to my knowledge no previous studies have considered motivational strategies found through analysis of a classroom corpus. Therefore, the first objective is to discover which strategies were found in the CLIL classrooms recorded and how these strategies differed depending on the use of AfL by the teacher over a range of academic subjects.

Part 1 (Chapter 6)	**Motivational L2 strategies in AfL and Non-AfL lessons**
Main objective: To compare discourse in AfL and Non-AfL classrooms with a focus on second language learning motivational strategies.	
RQ1	Do the frequency and distribution of second language motivational strategies differ depending on the use of AfL?
RQ2	How does the duration of these L2 motivational strategies vary depending on the subject (science, citizenship, art, drama)?
RQ3	Are there any L2 motivational strategies found in AfL lessons that are not identified in non-AfL lessons?
RQ4	Is there a relationship between teachers' use of AfL techniques and L2 motivational strategies observed during CLIL lessons?

The second part of the study involves collecting data from the students themselves regarding their motivation as well as feelings in classroom situations. This data was collected through questionnaires given to a small sample (N=40) of students at the end of the didactic unit.

Part 2 (Chapter 7)	**The effects of assessment on student self-reported motivation and feelings during the lesson**
Main objective: *To determine the effects of assessment techniques on student reported motivation and feelings in different classroom situations.*	
RQ5	Can any relation be seen between the type of assessment used by teachers and student's self-reported motivation? Are students in AfL classes more or less motivated than their non-AfL peers?
RQ6	How do AfL and non-AfL students describe their feelings in the context of certain classroom situations?

The third and final part of the study analyzes transcriptions of lower achieving student interviews using Martin and White's Appraisal framework (2005). The data for this part comes from interviews with a small sample (N=6) of lower achieving students from both AfL (n=3) and non-AfL (n=3) citizenship classes.

Part 3 (Chapter 8)	**Part 3: Lower achieving students' reflections on their own learning**
Main objective: *To determine the extent to which lower achieving learners express metacognitive reflection on their own learning through the use of various linguistic resources.*	
RQ7	How do lower achieving AfL and non-AfL students reflect on their own learning and classroom environment?

In sum, these research questions seek to determine if there is a relationship between the assessment techniques implemented in each classroom and the L2 motivational strategies identified in these CLIL lessons. These results are then compared to the motivation and the feelings of a sample of students (two groups), measured through self-reported motivation in questionnaires The study also considers a sample of lower achievers' reflections on their own learning and other factors related to their classroom environment.

5.3 The Corpus

The following sections give an overview of the participants involved and the data collection procedure. Data analyzed in this study comes from a larger corpus collected in bilingual primary schools during the 2010/2011 academic year.

5.3.1. Participants

Participants came from four primary schools in Madrid and its outlying areas. Each of these schools was fully integrated into the bilingual program, meaning the program spanned from Year 1 (5-6 years old) to Year 6 (11-12 years old) of primary education. Participation in the bilingual program stipulates that 40% of the curriculum is taught in English, and includes a variety of academic subjects, such as: literacy, science, art, drama, physical education and citizenship. Geographically, the locations of each school ranged from rural to suburban areas in Madrid, and the bilingual programs at each school were well established.

Student participants came from 5th and 6th grade classes and ranged in age from 10-12 years old. Each class was comprised of approximately 22 students, though certain lessons were taught as *desdobles*, meaning the teacher divided the class in half and gave separate instruction (on the same content) to each group. The students, with few exceptions, were native Spanish speakers, most of whom had been enrolled in the bilingual program since nursery school.

Table 5.1: Participants

School (teacher)	**Subject**	**Level (age)**	**Number of students (approx.)**
School 1 (AfL teacher 1)	Citizenship	Year 5 (10-11)	22
School 1 (AfL teacher 1)	Drama	Year 6 (11-12)	22
School 2 (AfL teacher 2)	Science	Year 6 (11-12)	22
School 3 (Non-AfL teacher 1)	Citizenship	Year 5 (10-11)	22
School 3 (Non-AfL teacher 1)	Art	Year 6 (11-12)	22
School 4 (Non-AfL teacher 2)	Science	Year 5 (10-11)	22
Totals **Four (4) schools/ teachers**	**Six (6) subjects**	**Years: 5/6** **Age range: 10-12**	**132 students (approximately)**

The four teachers involved in this project (three females, one male) ranged in age from mid-30's to mid-50's and had been working in bilingual education programs for at least seven years, though each had a minimum of 10 years of teaching

experience. The two AfL teachers were native English speakers from England and the United States, and the non-AfL teachers were Spanish nationals with a very high (C1/ C2) level of English. The AfL teachers had participated in a teacher-training seminar held by the British Council in which they learned AfL techniques which were then applied in their classes, though AfL Teacher 1 reported being trained in AfL techniques since the beginning of her teaching career in the UK. The two non-AfL teachers had no previous AfL training and were not familiar with methods of formative assessment.

5.3.2. Data collection procedure and corpus

Data collection took place over the course of an academic year (September 2010 to June 2011) and was conducted by myself and another doctoral student researcher from the Universidad Autónoma Madrid. The recordings form the primary school part of the UAM-CLIL Corpus, representing five bilingual schools throughout the Comunidad de Madrid. Of these schools, two were affiliated with the British Council while the remaining three belonged to the Comunidad de Madrid bilingual program. Access to the British Council bilingual schools involved in this study was facilitated by the British Council, which provided the initial contact information of teachers trained in AfL and administrators. Access to the remaining schools was based on the personal connections of the researchers involved, who contacted various teachers and administrators interested in collaborating with the investigation.

First, the teachers were contacted and asked if they would be interested in participating in the study. Upon confirmation, a meeting was set up with both teachers and administrators to describe the details of the project and procedures involved for recording lessons. A letter was drafted by two professors from the Department of English at the Universidad Autónoma de Madrid explaining the project to lend validity to the request for classroom access. In some cases, the school administrators requested additional approval from the Districto Area Territorial (DAT), a regional government entity that oversees educational administration for

each section of the city, which was requested and granted.[5] Before recording began, consent forms were sent home to the parents of each participating student describing the data collection process, and parents gave their consent for their child to appear in the recordings [Appendix 1].

The initial intention was to make a comparison of three AfL schools and three non-AfL schools. Access was given to record at two AfL schools and three non-AfL schools. Though attempts were made to contact a third AfL school in order to achieve an even distribution, requests to film were refused, first by the administration and later by the teacher. Due to time constraints and an already full schedule for filming, the decision was made to limit the recordings to five schools. Filming took place twice during the academic year: once at the beginning and once at the end, with the video recording of two separate didactic units in a variety of subjects. In some cases, more than one subject taught by the same teacher was recorded.

During the filming process, the teachers were asked to conduct their lessons with no special planning to achieve an authentic view of the discourse and activities taking place in each classroom. During the sessions, the camera stayed either off to the side of the class or moved around to observe students as they worked in groups. While at the beginning, some students were either shy or intrigued by the camera, they soon adapted to being recorded and continued with the class as if the camera were not present. The lessons lasted for an average of 45 minutes and included moments of the teacher conducting the lesson, student discussions, and group and pair work.

The video recordings from each of the five schools comprise a corpus of 82 lessons. After the recordings were transcribed, the word count totaled 475,999 words. The corpus was integrated into the UAM-CLIL project corpus representing primary schools, with full didactic units ranging from one to eight lessons in each subject. For

[5] Governmental permission is not necessary in Spain to research in primary schools unless requested by the head of studies. If the head of studies agrees to the terms of the study, this permission (along with that of the teachers/ parents) access for the researcher is granted.

the purposes of this dissertation, a selection of 14 classes was made, which will be further described in the next sections. Data from the student perspective was collected through a motivational questionnaire, administered to a sample of students (N=40) in citizenship classes after the end of their unit. Students from citizenship classes were chosen as the AfL citizenship teacher had more background with AfL training due to obtaining teacher education in the UK. In addition, interviews were conducted with three lower achieving students from AfL and non-AfL citizenship classes. These students were chosen for an interview by the teacher, and represent the bottom 15% of the class in terms of achievement (see table 5.2).

Table 5.2 Data description

Stage of Study	**Data Collection Method**	**Type of Data**	**Amount of Data**
1. L2 motivational strategies used by AfL and non-AfL teachers.	Video recordings of classroom interaction	Video recordings and transcripts	14 x 45 minute lessons (10 hours approx.) (71,504 words)
2. Student self-reported motivational questionnaire	Post-lesson administration of questionnaires to students from AfL and non-AfL citizenship lessons	Questionnaires	21 AfL 19 Non-AfL 40 total
3. Lower achieving student interviews	Interviews with three lower achieving students from each citizenship class	Video recordings and transcripts	6 x 7 min. (6,452 words)

With the overview of the data collection process and the corpus complete, sections 5.4-5.6 focus on the methods used for the three perspectives, which form the basis of this dissertation.

5.4 Motivational L2 strategies used by teachers in AfL and Non-AfL lessons (Part 1)

The first part of this study involves analyzing the corpus, identifying L2 motivational strategies used by teachers. The strategies were identified using an adapted version of the Motivated Orientation of Language Teaching (MOLT) scheme (Guilloteaux and Dörnyei, 2008). The dataset of classroom recordings is described first, followed by the instrument and analysis procedure.

5.4.1. Dataset: Corpus of classroom recordings

The 14 classes chosen for analysis represent six full didactic units in citizenship, science, art and drama classes. The units vary in length from one to four lessons. As mentioned previously, four teachers participated in the study. Two of the units were taught by the same teacher: AfL teacher 1 taught the citizenship and drama units while Non-AfL teacher 1 taught citizenship and art. Table 5.3 summarizes the corpus selected for analysis. The last column represents the total number of lessons in each unit and the duration of the unit.

Table 5.3: Corpus of classroom recordings

Teacher	**Subject**	**Unit Topic**	**School Year (age)**	**Assessment Type**	**Number of lessons/ duration**
AfL Teacher 1	Citizenship	Emotions	5 (10-11)	AfL	2 (1:21:14)
Non-AfL teacher 1	Citizenship	Democracy	5 (10-11)	Non-AfL	4 (2:41:01)

AfL Teacher 2	Science	Sound	6 (11-12)	AfL	3 (2:24:20)
Non-AfL teacher 2	Science	Prehistory	6 (11-12)	Non-AfL	2 (1:36:58)
AfL teacher 1	Drama	PET Speaking Practice, improvisation	6 (11-12)	AfL	1 (44:02)
Non-AfL Teacher 1	Art	Parallel and Perpendicular Lines	6 (11-12)	Non-AfL	2 (1:09:58)
Total	6 units			**Total**	14 lessons (9:57:33)

The data selection was based on finding AfL/ non-AfL parings in the same subjects (or closely related subjects, such as drama and art) to compare the strategies used during the lessons, avoiding as far as possible the influence of field or topic. Four academic subjects were recorded to collect a diverse cross section of data and to compare the use of motivational strategies depending on the content. It was hoped that with two teachers, the AfL training would outweigh the influence of teaching style.

While the lesson plan was not to be altered in any way due to the recording taking place, the AfL teachers were informed prior to the recording that the nature of the study was related to AfL. Both AfL teachers used a variety of techniques that reflect those described in Chapter 2: Assessment [see Appendix 2 for complete list].

5.4.2. Instrument: The MOLT Classroom Observation Scheme

The instrument used to measure the teacher's use of L2 motivational strategies was the Motivated Orientation of Language Teaching (MOLT) classroom observation scheme with some slight adaptations. As discussed in Chapter 3, the MOLT is based

on an adapted list of Dörnyei's 100 strategies that language teachers use to motivate students (2001), reduced to observable strategies. Measurement of these strategies is based on a real-time coding principle modeled in the communication orientation of language teaching (COLT) framework designed by Spada and Fröhlich (1995). The original classroom observation coding sheet for the MOLT scheme can be found in Appendix 3. While the original tested MOLT was used specifically in ESL lessons, in this case it was applied to CLIL lessons in which content was being delivered through the medium of a foreign language. Therefore, coding of the strategies was not limited to language-centered interactions, but rather encompassed content-based interactions, as well.

While the original MOLT framework contains 25 observable strategies for language teaching, I reduced the number to 15 strategies that were determined to have a relationship to AfL techniques. The reasoning for this reduction was based on the recommendation of a professor from the Universidad Autónoma de Madrid after a pilot study was completed on a small selection of lessons using 25 MOLT strategies related to teacher discourse. It was thought that the reduction of strategies would provide a more clear representation of any relationship found between the use of AfL and L2 motivational strategies. Additionally, the category of learners' motivated behavior (students' eager volunteering, engagement and attention) was discarded, as the principle aim of this part of the study was to measure the strategies in the teacher's discourse. Therefore, the number of strategies from the original MOLT scheme was reduced to 15 (see Table 5.4.). The relationship of each strategy to AfL is shown in the last column.

When completing a preliminary analysis, a trend developed wherein the teacher repeated a statement made by a student, which was originally coded as "neutral feedback". However, on closer inspection, including analysis of the intonation of the teacher, it was concluded that this type of feedback was actually being used as a way for the teacher to confirm the students' statement, and therefore could not be

considered neutral.[6] Therefore, an additional strategy called "echo" was added to the original MOLT scheme under the category of *encouraging positive retrospective self-evaluation*, making the total number 16. Table 5.4 shows the adapted version of the MOLT scheme, including a description of each strategy and the relationship to AfL.

Table 5.4: Adapted version of MOLT classroom observational scheme

1. Teacher Discourse		
MOLT Category 1.Referential Questions (Teacher discourse)	**Description** Asking questions to which the teacher does not already know the answer, including questions about the students' lives.	**Relationship to AfL** Effective questioning techniques
MOLT Category 2. Signposting (Teacher discourse)	**Description** Stating the lesson objectives explicitly or giving retrospective summaries of progress already made toward realizing the objectives.	**Relationship to AfL** Establishing clear learning objectives
MOLT Category 3. Stating the communicative purpose or utility of an activity (Teacher discourse)	**Description** While presenting an activity, mentioning its communicative purpose, its usefulness outside the classroom, its cross-curricular utility, or the way it fits into the sequence of activities planned for the lesson.	**Relationship to AfL** Establishing clear learning objectives
MOLT Category 4.Establishing relevance	**Description** Connecting what has to be learned to the students' everyday lives (e.g., giving grammatical examples with references to pop stars).	**Relationship to AfL** Promoting student engagement. Engaging students in the activity by allowing them to see how it is connected to them personally

[6] Thanks to Dr. Roy Lyster from McGill University in Canada for providing this valuable insight.

(Teacher discourse)		
MOLT Category 5. Arousing curiosity or attention (Teacher discourse)	**Description** During the presentation of an activity, raising the students' expectations that the upcoming activity is going to be interesting and/or important (e.g., by asking them to guess what they are going to do next, or by pointing out fun, challenging, or important aspects of the activity or contents to be learned).	**Relationship to AfL** Promoting student engagement. Engaging students in an activity from the beginning, drawing students into the class activity
MOLT Category 6. Scaffolding (Teacher discourse)	**Description** Providing appropriate strategies and/or models to help students complete an activity successfully.	**Relationship to AfL** Promoting student engagement. Teacher acting as a mediator to help students achieve goals/ close gaps
MOLT Category 7.Promoting cooperation (Teacher discourse)	**Description** Setting up a cooperative learning activity, or explicitly encouraging students to help one another, offering suggestions on how best to do this.	**Relationship to AfL** Promoting student engagement. Encouraging students to work together and giving them the opportunity to see exemplar work (i.e. lower achievers working with higher achievers)
MOLT Category 8.Promoting autonomy (Teacher discourse)	**Description** Offering students a choice of activities, involving them in making decisions regarding the timing of an activity, having them use the Internet or do research (e.g., for oral presentations, projects, and displays).	**Relationship to AfL** Promoting student engagement. Helping students take responsibility for their own learning
2. Participation Structure		
MOLT Category 9. Group Work (Participation structure)	**Description** The students are mingling, working in fluid pairs, or working in groups (simultaneously or presenting to the whole class).	**Relationship to AfL** Promoting collaboration. Encouraging students to work together and giving them the opportunity to see exemplar work (i.e. lower achievers working with higher achievers). Also gives the opportunity for peer and self-correction

MOLT Category 10. Pair work (Participation structure)	**Description** The students are working in fixed pairs (simultaneously or presenting to the whole class).	**Relationship to AfL** Promoting collaboration. Encouraging students to work together and giving them the opportunity to see exemplar work (i.e. lower achievers working with higher achievers). Also gives the opportunity for peer and self-correction
3. Activity Design		
MOLT Category 11. Personalization (Activity design)	**Description** Creating opportunities for students to express personal meanings (e.g., experiences, feelings, opinions).	**Relationship to AfL** Promoting student engagement. Engaging students in the lesson by giving them opportunities to express their own personal meaning.
4. Encouraging positive retrospective self-evaluation		
MOLT Category 12. Neutral Feedback (Encouraging positive retrospective self-evaluation)	**Description** Going over the answers of an exercise with the class without communicating any expression of irritation or personal criticism.	**Relationship to AfL** Feedback
MOLT Category 13. Process Feedback (Encouraging	**Description** Focusing on what can be learned from the mistakes that have been made, and from the process of producing the correct answer	**Relationship to AfL** Feedback

positive retrospective self-evaluation)		
MOLT Category 14. Elicitation of peer or self correction (Encouraging positive retrospective self-evaluation)	**Description** Encouraging students to correct their own mistakes, revise their own work, or review/correct their peers' work.	**Relationship to AfL** Peer and self-assessment
MOLT Category 15. Effective Praise (Encouraging positive retrospective self-evaluation)	**Description** Offering praise for effort or achievement that is sincere, specific (i.e., more than merely saying "Good job!"), and commensurate with the student's achievement. *N.B.:* Ability feedback ("You are very good at English") or praise involving social comparison ("You did better than anyone else in the class") is not recorded as *effective praise.*	**Relationship to AfL** Feedback
MOLT Category 16. Echo (Encouraging positive retrospective self-evaluation)	**Description** The teacher repeats what the student has said as a way of confirming their statement	**Relationship to AfL** Feedback

Adapted from Guilloteaux and Dörnyei, 2008

The process for analyzing the corpus of lessons according to these strategies is described in section 5.4.3.

5.4.3. Analysis procedure: MOLT

From the onset of this investigation, it was clear that the analysis procedure from the Guilloteaux and Dörnyei study in which the MOLT was first tested would have to be adapted. In the original study, researchers observed numerous lessons and used

a coding sheet to mark which L2 motivational strategies were put into practice by the teacher. The coding was completed on a minute- to- minute basis: as each minute of class elapsed, the researcher marked what had taken place during the previous minute (Guilloteaux and Dörnyei, 2008) and a tally of the total determined the duration of time for each strategy. The present study was to be based on a recorded corpus. Motivational strategies were identified in the transcriptions and timed from the video recording. This method made it possible to check and re-check the recordings and accurately record the duration of each strategy.

After lessons were recorded, they were transcribed and the transcriptions were uploaded into the UAM CorpusTool (O'Donnell, 2008), a software system that allows researchers to upload or create a scheme to code a corpus of written texts or oral interactions. Using the 16 strategies selected from the MOLT framework (see section 5.4.2), a scheme was created and then used to code the transcriptions, identifying L2 motivational strategies used by each teacher to determine the frequency and distribution of each strategy.

As the original MOLT scheme was based on a real time sampling, after coding the transcriptions, the video recordings of each lesson were used in order to measure the duration of each strategy. With the coded transcripts as a guide and using a stopwatch, the length of time devoted to each strategy in the recordings was measured to determine the length of time devoted to each strategy. Due to the time-sampling nature of this analysis, no overlapping strategies were coded. For example, if a teacher asked a referential question while stating the communicative purpose of the lesson, the overriding strategy (stating the communicative purpose) was coded to ensure that the times for each strategy were as accurate as possible. A time coding sheet was created with the length of time devoted to each strategy observed, as well as a brief summary of the lesson content [see Appendix 4 for sample time sheet].

After the first analysis was completed, a second analysis of the entire corpus was conducted six months later to determine the degree intra-rater reliability. The

recordings were reviewed a second time to ensure that the strategies originally identified were accurate and coincided with the definitions specified in the MOLT scheme. After making any necessary changes, the codings and times were adjusted [see Appendix 4 for sample re-coding sheet]. Additionally, in an analysis session with a group of graduate students of Applied Linguistics at the Universidad Complutense, several classroom extracts were coded by different raters, with a high degree of agreement.

Once this process was complete, the UAM CorpusTool was used to generate statistical data comparing the frequency and distribution of L2 motivational strategies found in AfL and non-AfL lessons. As the UAM CorpusTool did not have the capability to measure time increments, the real time codings were calculated manually to find the total duration, mean and range of each strategy (in minutes or seconds). These data were used to compare the amount of time devoted to L2 motivational strategies used in AfL and non-AfL lessons according to subject.

After the quantitative analysis was performed, extracts representing each of the 16 strategies were identified and analyzed qualitatively, taking into account the differences in usage of strategies by the AfL and non-AfL teachers, and if their usage had any relationship to AfL techniques. Thus, the data analysis represents the frequency, distribution and duration of L2 motivational strategies and continues with a qualitative analysis of extracts from the classroom transcriptions.

5.5 The effects of assessment on student self-reported motivation and feelings during the lesson (Part 2)

The second part of this study focuses on the student perspective, measuring learners' self-reported L2 motivation and feelings in certain classroom situations. Questionnaire data was collected from a sample of students in AfL and non-AfL citizenship lessons and analyzed to determine if any relationship could be found between the students' self-reported motivation and the number of L2 motivational techniques used by their teachers.

5.5.1 Dataset: Questionnaires

After the recording of citizenship units, the students (N=40)[7] were given a short questionnaire containing 20 items. The questionnaire contains two parts: the first was based on the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich and DeGroot, 1989) [Appendix 5] and the second was based on a metacognitive template which prompts students to describe their feelings in different classroom situations (Wall, 2008). Since the participants were in primary school, certain adaptations were made, as described in the next section. The questionnaire was administered at the end of the citizenship unit in both AfL and non-AfL classes. Students were instructed to be honest with their answers and told that their teacher would not have access to their responses. During the process and shortly after, some students expressed concern that the questionnaire was an exam, and were assured that this was not the case. Overall, the majority of participants seemed enthusiastic to participate.

5.5.2 Instrument: General student questionnaire: Adaptation of the Motivated Strategies for Learning Questionnaire (MSLQ)

The first two parts of the questionnaire given to citizenship students were adapted from the Motivated Strategies for Learning Questionnaire (MSLQ), which measures student motivational beliefs and self-regulated learning. The original questionnaire was used for a study conducted on college undergraduate students in the United States conducted by Pintrich and DeGroot (1989). Containing 56 items, it is divided into five motivational categories, which are listed in table 5.5 along with sample items.

[7] AfL: (n=19), non-AfL (n=21)

Table 5.5: Motivational categories and sample questions for MSLQ

Motivational Category	**Sample Items**
Self-efficacy	-I expect to do well in the course -My study skills are excellent
Intrinsic value	-I like what I am learning in the class
Test anxiety	-I get very nervous when I take an examination
Cognitive strategies	-When reading, I try to connect the ideas to things that I already know -When I study, I put important ideas into my own words
Self-regulation	-Even when the materials are uninteresting, I work hard until I finish -I work hard to get a good grade even when I don't like the class.

(Pintrich and DeGroot, 1989)

Since the items on the questionnaire were intended for university students, they were adapted be more accessible for primary school learners. For example, the scale from the original MSLQ, based on a seven point Likert scale (1=not at all true of me; 7=very true of me), was adapted to a more simplified scale based on three possible responses ("Yes" "No" and "I don't know" or "Yes", "No" and "Sometimes", depending on the section). The questionnaire was translated to Spanish to ensure participants had a full understanding of the items. A professor of psychology from the Universidad Autónoma de Madrid was consulted to ensure that both the translation and the adaptation of the questionnaire was acceptable and still contained the

elements of validity from the original MSLQ.[8] The adapted questionnaire can be found in Appendix 5.

5.5.2.a. Questionnaire: Part 1a

Part 1a of the questionnaire was designed for students to comment on the learning strategies used during the unit. The focus for this section was based only on the citizenship unit that had just ended to give students a chance to immediately reflect on their motivation. There were four items on this part of the questionnaire, and the scale used to measure the students' responses was "Yes", "Sometimes" and "No". The table below contains the English translation of two sample items from this part of the questionnaire.

Excerpt from Student Motivational Survey Part 1a

During the last unit….			
1. I raised my hand and participated	Yes	Sometimes	No
2. I asked for help when I didn't understand	Yes	Sometimes	No

5.5.2.b. Questionnaire: Part 1b

Part 1b of the questionnaire contained general questions about the subject, the semester and learning English. The goal was to measure the students' motivation and attitudes regarding the citizenship class and learning in English in general. The scale in this portion was slightly adapted ("Yes", "No" and "I don't know") due to the wording of each item. There were nine items in this section, taken from the five categories of the MSLQ.

Excerpt from the student motivational survey Part 1b

1. I liked the theme of the class	Yes	No	I don't know
2. I think that I did all of the homework well	Yes	No	I don't know
3. I understand as much as my classmates	Yes	No	I don't know

[8] Thanks to Dr. Héctor Grad for his expertise and assistance in adapting and translating the questionnaire.

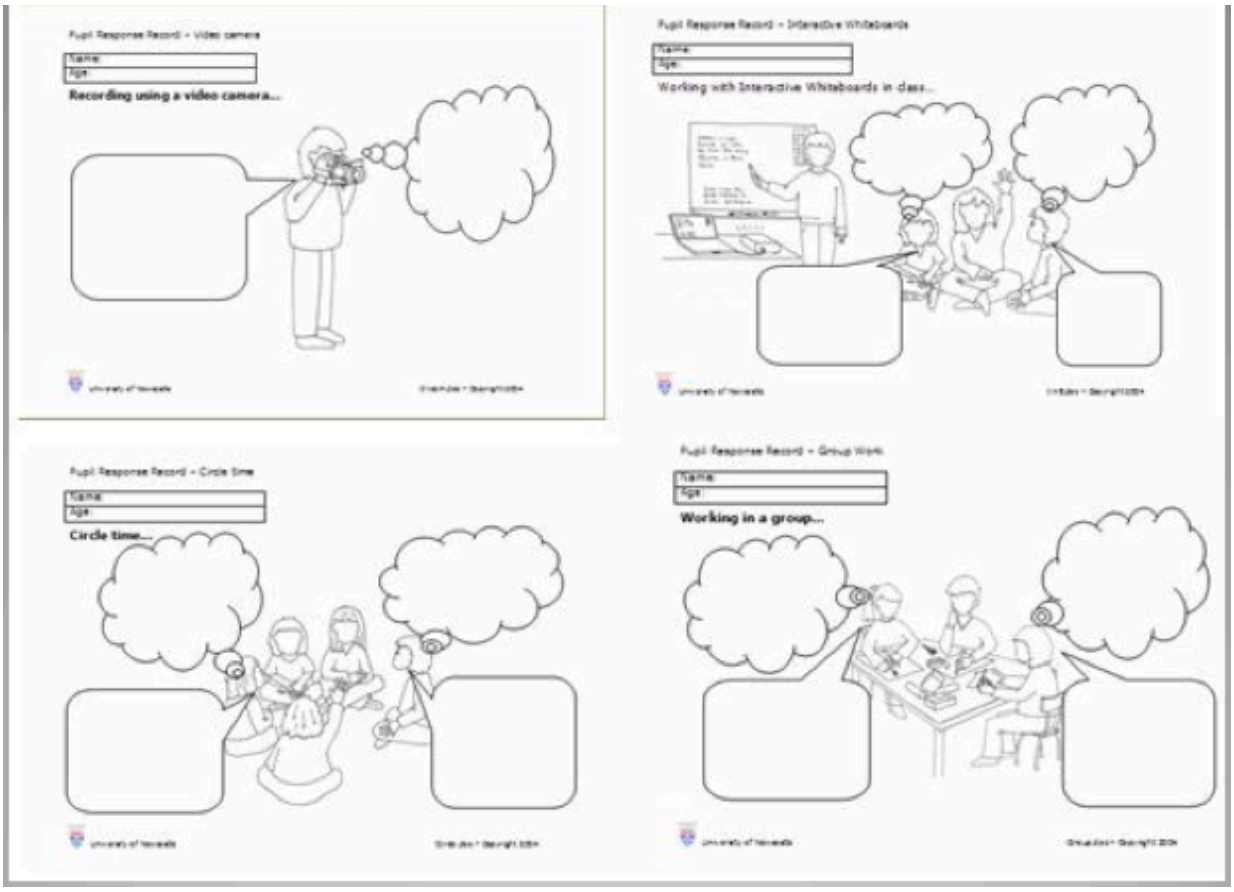
In sum, the aim for Parts 1a and 1b of the questionnaire was to document the strategies students were using throughout the unit as well as their motivation about the course and about learning in English in general.

5.5.3. Instrument: Student metacognitive templates (adapted)

The second part of the questionnaire was adapted from a metacognitive tool used to measure students' feelings during certain stages of the class, developed by assessment researcher Kate Wall (2008). The term metacognition refers to a learner's awareness of their own cognition and their ability to express this knowledge (Flavell, 1979). Rooted in psychological or semiotic tools mentioned by Vgotsky (1978), the templates are designed to give students a means of expressing themselves. Verbalizing their thoughts and emotions in different class settings gives students the ability to reflect on internal processes that teachers and researchers cannot access to better understand the learner's experience.

In the original study, students were given templates with pictures of classroom situations (e.g. circle time, working with the interactive whiteboard, working in groups). Above the drawing, a blank thought bubble and a speech bubble prompt students to report how they feel and what they say in these situations (Wall, 2008) (see figure 5.1).

Figure 5.1 Metacognitive templates for young learners



The combined use of speech and thought bubbles was meant to elicit information regarding both external and internal thought processes, providing students with way to express their beliefs on classroom dynamics, content and teaching. These templates have served as an empirical tool allowing researchers to gain insight into students' perspectives on learning and also as a pedagogical tool benefitting teachers (Wall, 2008).

Thus, Part 2 of the questionnaire was designed to reflect students' internal responses to classroom situations in an attempt to understand their feelings during different phases of the class. As discussed in Chapter 2, the goal of the AfL teacher is to provide a learning environment that promotes the emotional and psychological health of the students. Therefore, this part of the questionnaire was designed to

measure how comfortable students felt in the learning environment and their level of emotional well-being.

For the purpose of this study, the template was adapted: instead of being presented with a picture and a blank thought or speech bubble, students were provided with several different classroom situations and given adjectives to describe how they usually feel in each type of situation. They were provided with positive adjectives (e.g. happy, confident and intelligent) and negative adjectives (e.g. sad, shy, nervous and frustrated). The students were instructed to write one adjective for each situation and told that they could repeat the same adjectives for different situations if necessary. Some of the situations included:

Excerpt from student motivational survey Part 2

10. When I raise my hand in class I feel _____
12. When my classmates helps me I feel _____
13. When I don't understand something I feel _____
14. When I take an exam I feel _____

In order to make the students feel more comfortable in being honest with their responses, the questionnaire was translated to Spanish and checked by a native Spanish speaker from the UAM Department of Psychology for accuracy, giving the students the ability to respond in their native language. [See Appendix 5 for full questionnaire].

5.5.4. Analysis procedure: student motivational questionnaires

At the end of the unit after students had completed the questionnaire, the results of each part were compiled in graph form comparing the AfL and non-AfL groups. In order to analyze Parts 1a and 1b of the student questionnaires, each response given

by students was given a numerical value. Based on the nature of the responses, it was decided that "yes" was to be given a positive value (1), "sometimes/ I don't know" a neutral value (0) and "no" a negative value (-1)[9].

Table 5.6: Numerical values assigned to responses in Parts 1a/1b

Response	**Value**
Yes	1
Sometimes/ I don't know	0
No	-1

Once the values were assigned to each response, the means of the questions in each of the five categories were calculated using Microsoft Excel to compare the motivation of the two groups. These results provided a measurement of the self-reported motivation and the self-regulation strategies of each group of students. For a more detailed view of the responses, a graph was generated for each item based on the percentage of students who chose each response.

For Part 2 of the questionnaire, the results were compiled into bar graphs according to the number of students that selected each adjective. The graphs were then analyzed, comparing the responses of students from the AfL and non-AfL classes.

5.6 Lower achieving students' reflections on their own learning (Part 3)

The third and final part of this study seeks to analyze the ways in which lower achieving students are able to metacognitively reflect on their own learning based on post-lesson interview data.

[9] Thanks to Dr. Christiane Dalton-Puffer for her help with the scaling

5.6.1. Dataset: Lower achieving student interviews

For this part of the study, interviews conducted with three lower-achieving students from each of the AfL and non-AfL citizenship classes (N=6) were analyzed. These students were chosen by the teacher based on their performance during the academic year and represented the bottom 15% of the class. The students were asked to respond to open-ended questions to encourage them to reflect at length on their general motivation and experience in class after the completion of their citizenship unit. Each student was interviewed individually outside of the classroom to encourage them to answer openly and honestly. Due to the difficulty these students had expressing themselves in English, the interviews were conducted in Spanish.

5.6.2 Instrument: Lower Achieving Student Interview Questions: Adaptation of the Student Motivational State Questionnaire

Interview questions for lower achieving students were adapted from the Student Motivational State Questionnaire [see Appendix 6], the instrument used by Guilloteaux and Dörnyei (2008) in conjunction with the MOLT scheme in the study of ESOL students in South Korea. The items represent three categories: *attitudes toward the course; linguistic self-confidence and classroom anxiety*. A forth category was also added, *test anxiety,* to measure the students' attitudes toward this type of assessment. Rather than scaling the items, as was the case in the original study, the questionnaire was adapted and students were asked open-ended questions regarding their motivation in class.

Table 5.7: Sample items from Student Motivational State Questionnaire

Category	**Sample items**
Attitudes Toward the Course	-I like citizenship lessons this semester -When the class finishes I wish it would go on -I like my teacher
Linguistic self-confidence	-I feel like I am making progress in English this semester -I volunteer in class -I am sure one day I will be able to speak excellent English
Classroom anxiety	-I get very worried when I make mistakes in English -I worry that my classmates will laugh at me when I'm speaking English
Test anxiety	-I get more nervous when taking an exam in English than in Spanish

Adapted from Guilloteaux and Dörnyei, 2008

Since the items were originally designed for high school students, they were re-worded to make the content more accessible to younger learners. Several researchers and professors from the Universidad Autónoma de Madrid with an understanding of motivational questionnaires and second language research were consulted during the adaptation. The student interview questions can be found in Appendix 6.

5.6.3. Instrument: The Martin and White APPRAISAL scheme

The instrument used to analyze the lower achieving student interviews was the APPRAISAL scheme designed by Martin and White (2005). As described in Chapter 4, the scheme contains several sections and sub-sections to classify how stances,

opinions and beliefs are conveyed. In this case, two sub-sections of the category ATTITUDE were selected: JUDGEMENT and APPRECIATION. Both contained several sub-sections. An additional category, UNCLEAR, was added to classify instances of JUDGEMENT that did not clearly fit into one of Martin and White's categories. The coding process further classified each example as positive, negative or neutral to facilitate the comparison of the students' responses. Table 5.8 shows the reduced scheme used for this study [for full version including glosses see Appendix 7].

Table 5.8: Adapted Martin and White Appraisal framework: JUDGEMENT and APPRECIATION

Category	***Second classification***	***Third classification***	***Fourth classification***
Judgement	*normality*	*inscribed/ invoked*	*positive/ negative/ neutral*
	capacity		
	tenacity		
	veracity		
	unclear		
Appreciation	*reaction*	*impact*	
		quality	
	composition	*balance*	
		complexity	
	social valuation		

The next section describes the procedure used to analyze the interviews.

5.6.4. Analysis procedure

After the interviews were completed and transcribed, the transcripts were uploaded into the UAM CorpusTool and analyzed using the adaptation of Martin and White's framework (2005). While the framework provides specific categories and comprehensive definitions for each, the process of coding, especially for oral

communication, can be challenging, as other researchers have found (Read, Hope and Carroll, 2007; McCabe and Whittaker, forthcoming) Therefore, after a preliminary analysis, the transcripts were reviewed by a professor familiar with APPRAISAL theory and the more difficult items were negotiated. When this process was complete, statistical data of the frequency of each item was generated using the UAM CorpusTool. This was followed by a qualitative analysis of the extracts taken from interview transcripts. The analysis of the data aims to show how these lower achieving learners are able to self-assess and describe their learning process through the use of APPRAISAL.

5.7 Chapter summary

The purpose of Chapter 5 was to describe the datasets, instruments and analysis procedures used in this dissertation beginning with the research questions, which form the basis for the investigation. An overall view of the study was given, which included: a description of the corpus collected during the 2010/2011 school year; teacher and student participants from the four schools; and a description of the methods for data collection. The three perspectives of the study were then presented, beginning with "Motivational L2 strategies in AfL and Non-AfL lessons", which described the corpus of classroom recordings collected, the instrument used to analyze the corpus (the MOLT scheme) and the analysis procedure.

The chapter then discussed the methodology for the second perspective, "The effects of assessment on student self-reported motivation and feelings during the lesson", outlining the data collection procedure and the instrument. A review of the procedures for creating the questionnaire using an adaptation of the Motivated Strategies for Learning Questionnaire (MSLQ) and a metacognitive approach designed to measure students' emotions during certain classroom situations was given. The final part of the chapter presented "Part 3: Lower achieving students' reflections on their own learning", detailed the analysis procedure of the lower achieving student interviews using Martin and White's APPRAISAL framework (2005). Chapters 6, 7 and 8 are dedicated to presenting the results.

Chapter 6: Results: Motivational L2 strategies in AfL and non-AfL lessons

6.1 Introduction

This chapter presents results from analysis of the classroom corpus focusing on the first four research questions. The analysis explores the ways in which CLIL teachers make use of L2 motivational strategies during their lessons. This chapter compares strategies used by teachers with and without AfL training. The research questions addressed are the following:

Research Question 1: Do the frequency and distribution of second language motivational strategies differ depending on the use of AfL?

Research Question 2: How does the duration of these L2 motivational strategies vary depending on the subject (science, citizenship, art, drama)?

Research Question 3: Are there any L2 motivational strategies found in AfL lessons that are not identified in non-AfL lessons?

Research Question 4: Is there a relationship between teachers' use of AfL techniques and L2 motivation strategies observed during CLIL lessons?

As described in Chapter 5 (Methodology), the datasets for this part of the study consist of transcriptions from classroom recordings of six didactic units (subjects: citizenship, science, drama and art) analyzed using an adapted version of the MOLT scheme. The frequency, distribution and duration of each L2 motivational strategy are given, followed by classroom extracts analyzed qualitatively to show the relationship of these strategies to AfL practice. Through this analysis, a picture of how AfL and non-AfL teachers motivate their students begins to emerge.

The chapter begins with general statistics showing the frequency and distribution of each strategy found in the AfL and non-AfL corpus. Next, a real time analysis of the duration of each strategy is presented according to academic subject. Finally, the chapter gives a qualitative analysis of the 16 L2 motivational strategies found through the use of classroom extracts showing the relationship to AfL.

6.2 Overview of L2 motivational strategy findings

Table 6.1 presents L2 motivational strategies found in the AfL and non-AfL corpus. The first two columns show the frequency and distribution (%) of each strategy found in AfL lessons. The next columns show the frequency and distribution of strategies found in non-AfL lessons. The chi square shows the significance in the differences between the AfL and non-AfL distribution, and the significance of this chi square result is given in the last column.

Table 6.1 Overview of total L2 motivational strategies found in AfL and non-AfL lessons

	AfL Freq.	AfL Distribution	Non-AfL Freq.	Non-AfL Distribution	Chi Square	Signif.
Signposting	76	8.37%	67	8.87%	0.13	
Stating communicative purpose of activity	9	0.99%	1	0.13%	5.09	++
Referential questions	298	32.82%	363	48.08%	40.08	+++
Effective praise	77	8.48%	30	3.97%	13.91	+++
Neutral feedback	35	3.85%	51	6.75%	7.07	+++
Process feedback	6	0.66%	0	0.00%	5.01	++
Echo	104	11.45%	128	16.95%	10.39	+++
Elicitation of Self-or-peer-correction	48	5.29%	3	0.40%	33.15	+++
Arousing curiosity or attention	20	2.20%	3	0.40%	9.85	+++

Promoting cooperation	4	0.44%	5	0.66%	0.38	
Personalization	56	6.17%	20	2.65%	11.70	+++
Promoting autonomy	39	4.30%	10	1.32%	12.72	+++
Establishing relevance	23	2.53%	14	1.85%	0.87	
Scaffolding	77	8.48%	43	5.70%	4.77	++
Group work	12	1.32%	5	0.66%	1.77	
Pair work	10	1.10%	2	0.26%	4.03	++
Total frequency of strategies	**AfL** 908		**Non-AfL** 751			

Analysis of the table 6.1 reveals several significant differences between the L2 motivational strategies used in the AfL and the non-AfL lessons. First, the total number of strategies during the AfL lessons was higher (AfL=908, non=AfL=751). Differences in distribution are also evident: in AfL classes, strategies show a fairly even distribution while the majority of non-AfL strategies were found in the categories: referential questions (48.08%) and echo (16.95%). This difference in distribution is evidenced by the significant differences between the percentages in many categories, as many strategies in the non-AfL lessons were used infrequently. The distribution of strategies was higher in AfL lessons in 13 out of 16 of the motivational categories, with the exception of referential questions, neutral feedback and echo. Sections 6.5-6.10 elaborate on these findings with an analysis of classroom extracts found in each category.

6.3 Real time analysis of teacher L2 motivational strategies by subject

This section presents the results obtained through real time coding of each L2 motivational strategy, comparing the duration of each by subject. The duration of each L2 strategy was measured in real time in the video recordings using a stopwatch. Table 6.2 shows the total duration of time L2 motivational strategies

were used during each unit. Comparing this time to the total duration of each unit, the percentage of class time devoted to the strategies was calculated.

Table 6.2 Ratio of time devoted to L2 motivational strategies by subject

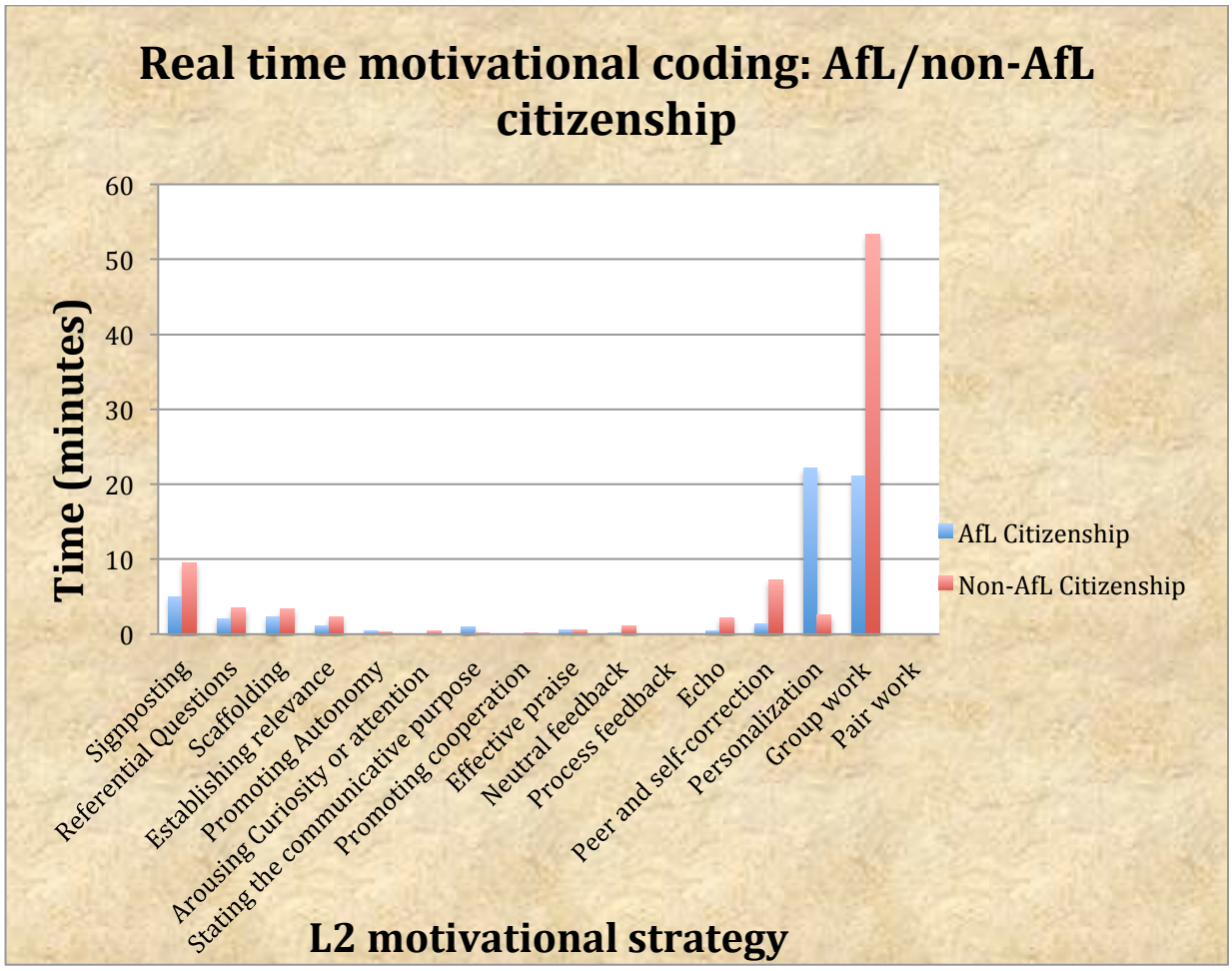
	Time devoted to L2 motivational strategies during unit	**Total duration of unit**	**Ratio of time L2 motivational strategies used compared to duration of unit**
AfL citizenship	57 minutes	1 hour 21 minutes	71%
Non-AfL citizenship	1 hour 16 minutes	2 hours 1 minute	54%
AfL Science	1 hour 10 minutes	2 hours 24 minutes	48%
Non-AfL Science	43 minutes	1 hour 36 minutes	43%
AfL Drama	21 minutes	44 minutes	48%
Non-AfL Art	9 minutes	1 hour 9 minutes	14%

In all three units, the percentage of time devoted to L2 motivational strategies was higher in the AfL units. The highest percentage of time devoted to motivational strategies was found in AfL citizenship lessons (71%). The percentage in the non-AfL citizenship unit was also quite high (54%). The strategies were used for a similar duration of time in AfL (48%) and non-AfL (43%) science lessons. In the AfL drama class, the ratio of time strategies were used was 48%, and the lowest duration was seen in the non-AfL art lesson, which accounted for only 14% of the total. For a table of the range, mean and duration of each strategy, see Appendix 8.

6.3.1 Differences in L2 strategy duration: citizenship

Figure 6.1 shows the duration of L2 motivational strategies found in AfL and non-AfL citizenship lessons. The scale (measured in minutes) reflects the longest time a strategy was used, in this case group work (53 minutes).

Figure 6.1: Real time motivational coding results: citizenship lessons

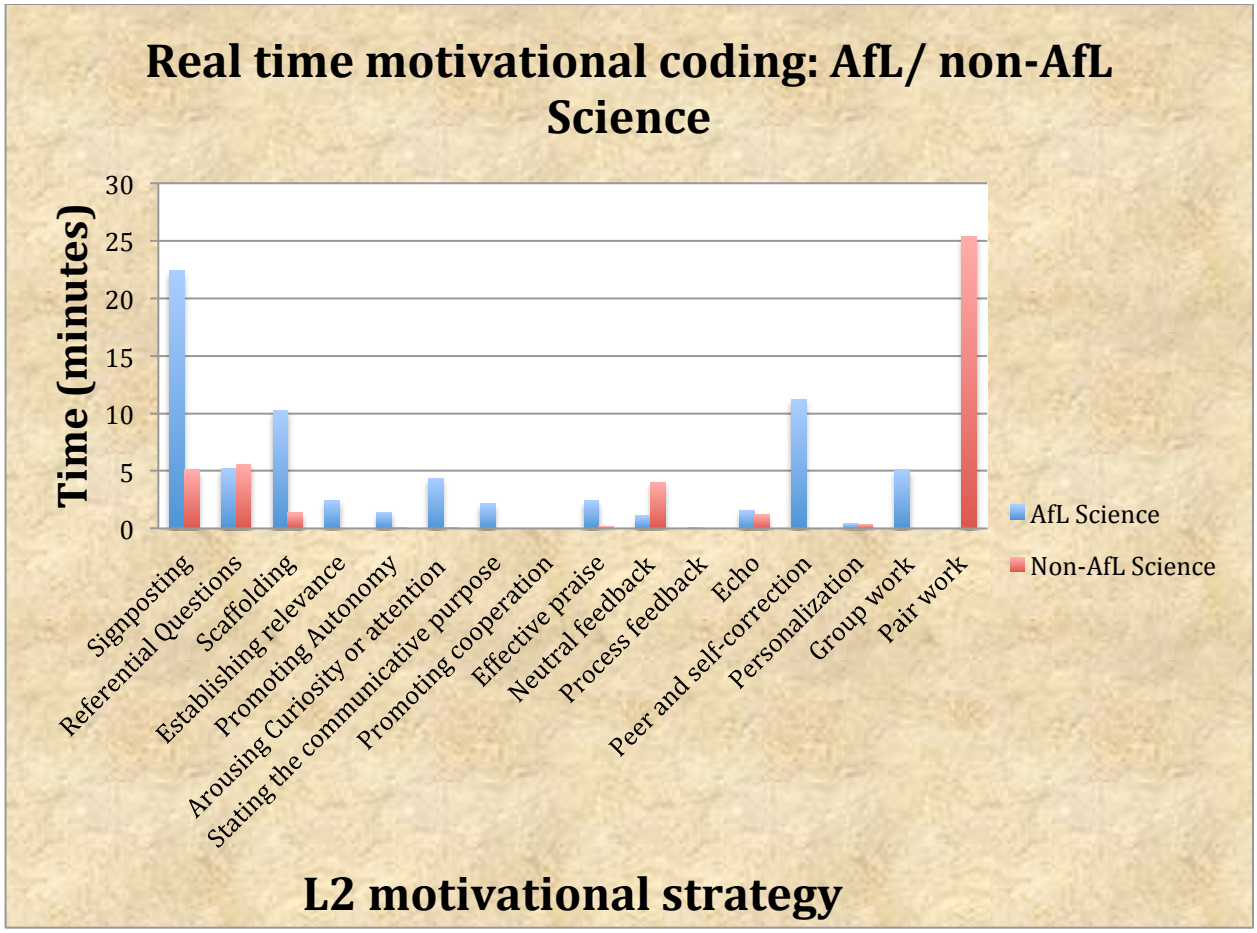


When comparing the citizenship lessons, the duration of motivational strategies is similar. However, differences are observed in specific categories, namely group work. The non-AfL students spent a large portion of the unit working in groups as they were engaged in a unit on democracy and conducting a debate. Peer and self-correction and signposting also occurred for a longer duration during Non-AfL lessons. Referential questions, scaffolding and establishing relevance were also used for a slightly longer duration, through this increased time was minimal. In the case of the AfL lessons, the only strategy used longer than in non-AfL lessons was personalization, as the students were completing a unit on emotions in which the teacher encouraged them to discuss personal experiences.

6.3.2 Differences in L2 strategy duration: science

Figure 6.2 shows the duration of time of each motivational strategy in AfL and non-AfL science lessons. In this case, the scale reaches 30 minutes.

Figure 6.2: Real time motivational coding results: science lessons



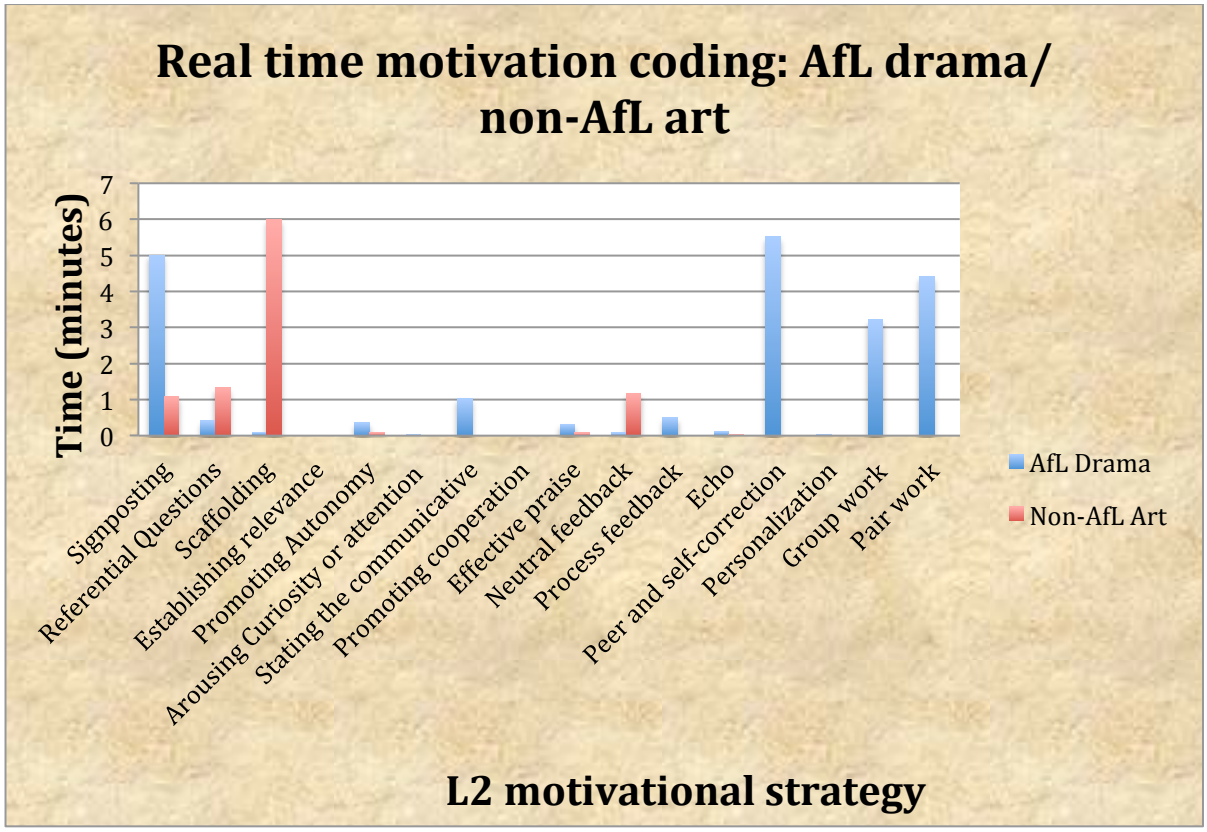
When considering the science units, a greater difference between the AfL and non-AfL groups is visible. In general, the AfL teacher used more motivational strategies for a longer duration. In this case, the total duration of ten motivational strategies was higher during the AfL unit (signposting, arousing curiosity and attention, stating the communicative purpose, effective praise, peer and self-correction and group work). There was an especially high duration (>20 minutes) of signposting during the AfL unit.

In the non-AfL lessons, only six strategies were found during science lessons, with the longest duration devoted to pair work. Signposting, referential questions, scaffolding and neutral feedback were also used.

6.3.3 Differences in L2 strategy duration: AfL drama and non-AfL art

Figure 6.3 shows the duration of L2 motivational strategies used in AfL drama and non-AfL art lessons. For these lessons, the total duration of strategies used was lower with the scale reaching only seven minutes.

Figure 6.3: Real time motivational coding results: AfL drama and non-AfL art lessons



Comparing the AfL drama and non-AfL art class, numerous differences are seen. Here, the duration of strategies was lower than those found in the citizenship and science lessons. Nevertheless, a variety of strategies were used in the AfL drama lesson.

In the AfL lessons, shorter instances of all strategies were found with the exception of establishing relevance. The strategies with the longest duration were signposting

and peer and self-correction. Instances of group and pair work were also found, as it was common for the AfL teacher to divide students into groups to practice speaking, improvise or rehearse scenes.

In the non-AfL art unit fewer strategies were found. Scaffolding was seen quite frequently, as the non-AfL teacher was active in assisting students as they completed their assignments throughout the classes.

6.4 Organization of individual L2 motivational strategy results

Sections 6.5-6.10 analyze each L2 motivational strategy found in AfL and non-AfL lessons. Instead of presenting the results according to the categories identified in the MOLT scheme, the results were classified according to AfL techniques (stating the lesson objectives, effective questioning techniques, feedback and peer and self-assessment). The last two categories (promoting student engagement and autonomy, promoting collaboration) represent two of the objectives of AfL also described in Chapter 2. The aim of this organization is to discover the possible relationship of these strategies with the teachers' use of AfL. The organization is shown in table 6.3.

Table 6.3: Organization of L2 motivational strategies results

AfL technique	**L2 Motivational strategies**
6.5 Establishing clear learning objectives	6.5.1.Signposting 6.5.2 Stating the communicative purpose
6.6 Effective questioning techniques	6.6.1. Referential questions
6.7. Feedback	6.7.1 Effective praise 6.7.2 Neutral feedback 6.7.3 Process feedback 6.7.4 Echo
6.8 Peer and self-assessment	6.8.1 Peer and self-Correction

6.9 Promoting student engagement and autonomy	6.9.1. Arousing curiosity or attention 6.9.2. Promoting cooperation 6.9.3 Personalization 6.9.4. Promoting autonomy 6.9.5 Establishing relevance 6.9.6 Scaffolding
6.10 Promoting collaboration	6.10.1 Group Work 6.10.2. Pair Work

Sections 6.5-6.10 present the results of L2 strategies found in the corpus of recorded CLIL lessons. Each section presents a description of each strategy followed by the frequency and distribution found during AfL and non-AfL units. The section continues with extracts from the classroom transcriptions and an analysis of how the strategies were used by AfL and non-AfL teachers.

6.5 AfL category: Establishing clear learning objectives

6.5 Establishing clear learning objectives	6.5.1.Signposting 6.5.2 Stating the communicative purpose

6.5.1. Signposting

Signposting is used predominately at the onset of the lesson as a means of stating the objectives of the day or reflecting on what has been done in previous lessons.

Signposting Stating the lesson objectives explicitly or giving retrospective summaries of progress	**AfL frequency** 76	**AfL distribution** 8.37%	**Non-AfL frequency** 67	**Non-AfL distribution** 8.87%	**Chi Square** 0.13	**Signif.**

Signposting represents 8.37% of AfL strategies (76) and 8.87% of non-AfL strategies (67), a similar distribution with no significant differences. One of the ways in which both groups of teachers used signposting was as a means of interacting with

students. This was usually done at the beginning of a lesson to introduce the learning objectives of the day and encouraging recollection of what had been learned in a previous lesson to tie the content together. Teachers trained in AfL often made use of the WALT (We are learning to…) and WILF (What I'm looking for) posters when engaging in this form of signposting, as seen in extract 6.1.

Extract 6.1 AfL Signposting: WALT and WILF

1 ***AfL TCH***[10]*: What have we been doing in citizenship? ((puts the WALT and WILF cards up on*
2 blackboard)) STU[11] *1?*

3 ***STU 1:*** *Emotions*

4 ***AfL TCH:*** *We were looking at emotions, right? We started looking at emotions. Who can tell*
5 me anything that we did in Wednesday's lesson about emotions? STU 2?

6 ***STU 2****: Em, in Wednesday, we did, you give us a paper that we have to color,*

7 ***AfL TCH:*** *And then?*

8 ***STU 2****: And we had to color it with the color that we think that emotions ….the*
9 personality in the color

In this interaction, the teacher asks students to reflect on what they had been doing in their citizenship class up to that point. The teacher incorporates WALT and WILF posters to remind students of the objectives for the class (Lines 1 and 2). After receiving a visual reminder, the student is able to respond with the theme of the lesson, after which the teacher elaborates by asking a follow-up question, prompting the class to recall and describe the specific activity (Lines 4-7). This type of signposting serves a duel purpose, as it engages students in the learning process by allowing them to display their knowledge and puts them in a position of authority as the "knower". The AfL technique of incorporating visual clues through the WALT is helpful in setting these learning criteria, as it provides students with a base for understanding the aims, which the teacher may refer back to in future lessons in order to give continuity. Extract 6.2 gives an example of signposting in the non-AfL classroom.

[10] TCH=Teacher
[11] STU=Student

Extract 6.2 Non-AfL signposting

1 ***Non-AfL TCH:*** *So every group is going to have one paragraph, you will read about it, you will*
2 read sorry, and you will mention the main ideas. We were speaking about democracy and
3 what did we do? You remember what we did last?

4 ***STU:*** *Em*

5 ***Non-AfL TCH****: Shh, wait, STU 1*

6 ***STU 1****: We talk about one paragraph*

7 ***Non-AfL TCH:*** *Uh huh*

8 ***STU 1:*** *And (inaudible)*

9 ***Non-AfL TCH:*** *Of the?*

10 ***STU 1****: Of the democracy.*

11 ***Non-AfL TCH:*** *Of democracy. Good, and we had the text about democracy. Very good. Ok,*
12 and what, were we working individually or in groups?

13 ***STU****: In groups.*

14 ***Non-AfL TCH****: All right, and we read the most important ideas of everybody*

The teacher begins the interaction by stating the task for the day and asking students to recall the activity from the previous lesson (Lines 1-3). The students struggle to remember the content of the previous lesson, and the teacher calls on Student 1 (Line 5) who recalls that they were focusing on one paragraph from an article. The teacher then prompts the student to remember the topic (Lines 9-10) and the participation structure of the previous lesson (Lines 11-13) and ends the interaction by elaborating on the information the students had provided (Line 14). Through this exchange, the teacher is able to establish the current task and also encourage students to reflect on what they had done in the previous lesson, though the students initially showed difficulty in remembering the theme of the lesson. Visual clues such as the WALT and WILF, as seen in the previous AfL extract, enabled students to immediately remember the theme of the unit.

Signposting plays a key role in informing students of what they will be learning from the beginning of the lesson. Making these aims clear enables students to make a checklist of the knowledge they are responsible for, which they can later return to in order to fill any gaps. Stating the learning aims from the beginning of the lesson is an important factor in guiding the learning process. In extract 6.3, the Non-AfL teacher explains what students are learning next.

Extract 6.3 Non-AfL stating the learning aims

Non-AfL TCH: *Ok, so we are gonna start with.. eh.. a few history units there are in the book, and the first unit is about Prehistory and pre-Roman times.*

The non-AfL teacher begins this social science lesson by establishing a focus on history and directing students to the first unit in the book with the topic of Prehistory and pre-Roman times. This statement allows students to anticipate what they will be learning for the next several lessons. However, the statement does not address the aims for learning, which could serve to enhance the students' expectations for the unit and outline what they are expected to know. In extract 6.4, an AfL teacher asks students to read the learning objectives, which are displayed on a poster in the front of the class.

Extract 6.4 AfL Stating the learning aims

1 ***AfL TCH****: What are you learning today? Please, since you are the students, could you read to* *2 each other? We are learning to understand…*
3 ***STU:*** *((reading)) how different phenomenon happens and what the different characteristics* *4 of sound are. To use texts to write answers to the questions correctly.*

In this interaction, the teacher asks students what they are learning today by displaying visual clues, asking them to read the objectives aloud as a group (Lines 1 and 2). As the students read, they becomes aware of the general aims that they must master (Lines 3 and 4) and learn what is expected of them during that lesson at the onset. Giving students clear aims demystifies the learning process by providing targets they must work toward. By highlighting the learning points visually,

students can achieve these targets in a more efficient way. Furthermore, explicitly outlining the learning aims accompanied by the requirements for achieving each grade gives students the tools to know where they stand and how they can reach the next level, as seen in extract 6.5.

Extract 6.5: AfL stating explicit learning aims

1 ***AFL TCH****: You need….we need to know if you have a Sufi, a Bien, Notable, a Sobre*[12] *or what*
2 So here around the room, there is a poster ((points)), there is a poster ((points)) there is a
3 poster ((pointing)) and here is a poster, ok? So, for the difficult things a Sufi is a Sufi.
4 ((reading aloud)) I can recognize some phenomena produced….produced ….produced by
5 sound. I can describe the pitch and the loudness. These words I don't understand! Is that ok?

6 ***STU****: Yes*

7 ***AfL TCH:*** *Yes, because we just started, so during the unit, you will have to understand these*
8 things for a Sufi. You will have to understand more things, because you have to explain some
9 phenomenon. Explain. That's more difficult. But we have two days to do it. And here in the
10 Notables: ((reading)) explain how sound moves through solids and liquids and gasses…

Through visual means, students are presented with a scheme of objectives, which the teacher uses to elaborate on the aims to achieve each grade (Lines 2-3). The teacher begins by reading the requirements to achieve the lowest grade (*sufi*), which represents the baseline criteria that students must master in order to pass (Lines 3-5, 7-9). Moving up to the higher grade, *notable*, the teacher reveals the higher order abilities (describing, explaining) that they are expected to master to achieve this grade (Line 10), noting the rise in the level of difficulty (Lines 8-9) while encouraging them by saying that mastery is possible (Line 9). This approach of using "I can statements" illustrates one of the main objectives of AfL: making the learning criteria clear by setting achievable goals.

6.5.2 Stating the communicative purpose

Stating the communicative purpose of the activity involves explaining the usefulness of an activity in the real world or explaining the link between the activity to another lesson or class.

[12] Grades in Spanish grading system *Sufi* (sufficient), *Bien* (good), *Notable* (notable), *Sobre* (excellent)

Stating communicative purpose of activity When presenting activity, mentioning its usefulness outside of the classroom, its cross-curricular activity or the way it fits into a sequence of events planned for the lesson	**AfL frequency** 9	**AfL distribution** 0.99%	**Non-AfL frequency** 1	**Non-AfL distribution** 0.13%	**Chi square** 5.09	**Signif.** ++

This strategy was one of the most infrequent in both the AfL (9/ 0.99%) and the non-AfL (1/0.13%) classrooms, though significant differences were found between the two groups (chi square=5.09, $p<0.05$). While few instances appeared in the corpus, it is worthwhile to note the ways the purpose of an activity was conveyed mainly during citizenship lessons.

In extract 6.6, a non-AfL teacher explains the purpose of the lesson as the class begins a debate. This was the only instance of this strategy in the non-AfL lessons.

Extract 6.6 Non-AfL Stating the communicative purpose of a debate

1 ***Non-AfL TCH****: Now we are going to have a debate about wearing a uniform, and you know*

2 that in a debate, what do we do? When we are debating? STU 1?

3 ***STU 1****: We, we, um, we say our opinions*

4 ***Non-AFL TCH****: Ok, you have to express your opinions*

The teacher begins by providing an explanation of the objectives for the lesson (Lines 1-2). The nature of this activity is both practical and relevant, as the students are learning about democracy and given the opportunity to mimic a democratic process, debating on a subject that is especially relevant to them. The students are asked about the purpose of a debate (Lines 3-4), concluding that it is an opportunity to express their opinions. While the teacher does not explicitly describe the utility of a debate outside of the classroom, this activity is developed to engage higher order thinking skills by being given a forum in which students state their opinions and listen and evaluate the opinions of their classmates.

In extract 6.7, the AfL citizenship teacher explains the learning objectives, which focus on identifying different types of emotions.

Extract 6.7: AfL Stating the communicative purpose: emotions

1 ***AfL TCH****: We are learning to understand that there are different types of emotions and*
2 different ways to express them. We are learning to understand my emotions, the triggers,
3 which means the things that make you feel these emotions, and ways to deal with different
4 emotions, ok? And when we know we can do that is ((reading from WILF poster)) I can name
5 different emotions, I can tell you when I feel that emotion, reasons why I may feel that way
6 and how I react.

When stating the objectives for the day through the use of the WILF poster, the teacher explains that the lesson focuses on different types of emotions. However, the aims also explicitly deal with the ways these emotions are triggered in everyday life, and strategies for expression (Lines 1-3). While the topic itself relates to daily life, the teacher nevertheless uses the objectives to explain the points that the lesson will cover (Lines 4-6) and how they might be applied outside of the classroom. At the end of the dialogue, the teacher re-states the learning aims using a combination of "I can statements" and WILF, specific AfL tools (Lines 5-6).

6.5.3 Summary of establishing clear learning objectives

The findings show that signposting was found with the same frequency in the AfL and non-AfL lessons, while stating the communicative purpose was found more in AfL lessons. Based on a qualitative analysis of the extracts, the incorporation of mediating artifacts, such as WALT, WILF and "I can" statements serve to enhance these motivational techniques, as they are designed to provide students with visual clues and explicitly stated aims (see extracts 6.1, 6.5 and 6.7). These artifacts also facilitate continuity, as the teacher may refer to them as the unit progresses, enhancing student memory of the material that has been covered in previous lessons.

6.6 AfL Category: Effective questioning techniques

6.6 Effective questioning techniques	6.6.1. Referential questions

6.6.1. Referential questions

The most frequently used strategy by AfL and non-AfL teachers was the use of referential questions, or asking questions to which the teacher does not know the answer.

Referential questions Asking the class questions to which the teacher does not know the answer	**AfL frequency** 298	**AfL distribution** 32.82%	**Non-AfL frequency** 363	**Non-AfL distribution** 48.08%	**Chi square** 40.08	**Signif.** +++

During the units, the AfL teacher asked a total of 298 referential questions, which comprised 32.82% of the total motivational strategies used during the units. The non-AfL teacher asked a total of 363 referential questions, which accounted for 48.08% of the total motivational strategies. The difference was highly significant (chi square=40.08, $p<0.02$). Use of referential questions can be further classified into different types depending on the purpose, making a distinction between the regulative versus instructional register. The regulative register refers to interactions related to classroom organization or management, while the instructional register is related to the transmission of content (Christie, 1995). While both types of questions were included in the final count, the extracts selected focus only on questions relating to the instructional register.

All teachers asked open referential questions (questions which require extended answers from the student) and closed referential questions, (questions which request a simple one-word, short or yes/ no answer) related to the content. Extract 6.8 presents examples of closed referential questions.

Extract 6.8: Closed referential questions

Example 1: ***Non-AfL TCH:*** *You understand what freedom is?*

Example 2: ***Non-AfL TCH:*** *Do you think it´s a straight line?*

Example 3: ***Non-AfL TCH:*** *Did you answer that it was in Athens?*

Example 4: ***Non-AfL TCH:*** *What are you going to vote for?*

Example 5: ***Non-AfL TCH:*** *Does any of you know the difference between prehistory and history?*

****

Example 6: ***AfL TCH:*** *Do you like this symbol better than that symbol?*

Example 7: ***AfL TCH:*** *Was that easy or hard?*

Closed referential questions related to content were more frequent in non-AfL lessons. The main purpose of the questions aimed to measure students' previous knowledge on specific subject matter, for example when the teacher inquires as to whether they know the difference between history and pre-history (Extract 6.8, Example 5). In terms of questioning techniques, the non-AfL teachers posed yes or no questions (Extract 6.8, Examples 2, 3, 5) while the AfL teachers gave students two options to choose from (Extract 6.8, Examples 6 and 7). In general, the use of closed referential questions was found more frequently in non-AfL classrooms than AfL classrooms.

Open referential questions related to lesson content were found frequently in both AfL and non-AfL classrooms, encouraging students to provide an explanation or description in their response.

Extract 6.9: Open referential questions

Example 1: ***Non-AfL:*** *Why should we wear a uniform or why shouldn´t we wear a uniform?*

Example 2: ***Non-AfL:*** *Can you explain to me what a debate is?*

Example 3*: Non-AfL:* *How do you think we know all these things about people that lived in prehistory?*

Example 4*: Non-AfL:* *Can you tell me some differences between one period and another?*

****

Example 5: ***AfL TCH:*** *So what do we know about thunderstorms? What do we know?*

Example 6: ***AfL TCH:*** *What did you discover?*

Example 7: ***AfL TCH:*** *How could you get out all your anger without hurting anybody?*

Example 8: ***AfL TCH:*** *What reason might there be for her to be sad?*

Referential questions encouraged students to explain certain concepts further, such as in debates (Extract 6.9, Example 2). These questions also provided an opportunity for the teacher to measure students' previous knowledge on a subject (Extract 6.9, Example 4, Example 5) or to formulate an opinion (Extract 6.9, Example 1, Example 8). In all cases, the teachers are using questions to allow students to elaborate on their knowledge or opinions. Such exchanges give the teacher information, which may then be used when assessing the learners´ knowledge or ability level related to a certain concept. In the CLIL classroom, open referential questions are also a means of evaluating the learners´ ability to articulate a certain concept in the foreign language. Regardless of the purpose, asking open questions is an effective way of creating student output, which the teacher can then assess. The findings show that AfL and non-AfL teachers used open questions in an effective way.

A common feature found in the AfL classroom was asking follow-up questions. These questions clarified the students´ answer to the previous question, or ask for elaboration or additional information.

Extract 6.10 AfL referential questions: follow-up

Example 1: ***AfL TCH:*** *Can you say that in an easier way?*

Example 2: ***AfL TCH:*** *Can you explain that?*

Example 3: ***AfL TCH:*** *If I don´t understand, what should I do?*

In Extract 6.10, Examples 1 and 2 show the teacher requesting clarification, while in Example 3 the teacher uses a metacognitive strategy to encourage students to reflect on the learning process. By asking students to visualize a solution, the teacher helps them reflect on possible courses of action. This strategy then enhances

the ability for students to self-regulate their learning.

Evaluation-based questions, which encourage students to think about and evaluate their own learning objectives, were also found only in AfL classrooms. The purpose of this question type was to prompt learners to reflect on the objectives of the day.

Extract 6.11 AfL referential questions: self-reflection on learning objectives

Example 1: ***AfL TCH:*** *So what do you have to do?*
Example 2: ***AfL TCH:*** *What are you learning today?*
Example 3: ***AfL TCH:*** *What do you have to know?*
Example 4*: AfL TCH:* *Now, can you go to the next level?*

In Examples 1 and 2 of Extract 6.11, the questions are presented at the beginning of the lesson as a part of the WALT and WILF, give the daily learning objectives. This technique has the same purpose as *signposting*, differing in that the teacher does not reveal the aims, but rather encourages the students to think and decide their goals and tasks independently. Similarly, Example 3 of Extract 6.11 engages students in a reflection on the most important points they are expected to know. Finally, the last question, asked in the context of ¨I can¨ statements, refers to concrete objective levels with a detailed description of what the students are expected to know. The teacher encourages student reflection on their own learning, asking them to self-assess their ability level and whether they are ready to progress based on their current knowledge. By asking these types of questions, the teacher is opening up a forum for self-evaluation.

6.6.2 Summary of effective questioning techniques

Referential questions, while common in both AfL and non-AfL lessons, were used differently by each teacher and represented the highest frequency of L2 motivational strategies found in the corpus. Open and closed questions related to content were found in both groups, though the closed questions were found more frequently in the AfL lessons. The purpose of the questioning ranged from measuring students' previous knowledge to encouraging them to state opinions and elaborate on the material. Question types found only in the AfL lessons include

follow-up questions, or building upon previous answers by asking students for further explanation. Additionally, AfL teachers were more likely to use referential questions to encourage students to reflect on the learning objectives and self-assess their level.

6.7 AfL category: Feedback

6.7. Feedback	6.7.1 Effective praise 6.7.2 Neutral feedback 6.7.3 Process feedback 6.7.4 Echo

6.7.1 Effective praise

Effective praise offers positive feedback to students based on their achievement. The MOLT scheme specifies that this praise must go beyond general encouragement and instead must be specific.

Effective praise Offering praise for effort or achievement that is sincere, specific and commensurate with the student's achievement.	**AfL frequency** 77	**AfL distribution** 8.48%	**Non-AfL frequency** 30	**Non-AfL distribution** 3.97%	**Chi square** 13.91	**Signif.** +++

In the AfL units, a higher frequency of effective praise was found, with 77 instances (8.48%). In non-AfL lessons 30 instances were found (30/3.97%), with a highly significant chi square result (chi square= 13.91, p<0.02). Extract 6.12 shows examples of effective praise in AfL lessons.

Extract 6.12 AfL effective praise

Example 1
AfL TCH: *Stop a minute! They have an incredible discovery! Can you tell them, STU?*

Example 2
AfL TCH: *I see some students looking in their books for this information because you are science detectives and you have to find the information*

Example 3
AfL TCH: *Very good! You had to think of a color that matched that emotion, didn´t you? And you did a brilliant job!*

Example 4
AfL TCH: *Very logical, very nice, very good way of doing it*

Example 5
AfL TCH: *Because we are going up and up and up*[13]

In Examples 1 and 2 of Extract 6.12, taken from a science lesson, the AfL teacher puts students in the role of ¨science detectives¨, or active learners, asking them to research a specific scientific phenomenon. As the students present their results, the teacher praises one group on a discovery that they have made, asking them to share it with the class. This situation is an example of the AfL approach of "teacher as mediator" encouraging students to complete their own investigation while offering support and feedback. As the investigation continues, the teacher then points out the students that are researching the way that they should, commending them on being science detectives and affirming their work as they find the information they are seeking.

In Example 3 of Extract 6.12, the AfL citizenship teacher praises the class as a whole for completing a task, which involves finding a color for a specific emotion. Later, the teacher gives more praise after one student has given their explanation, noting the logic behind the statement (Extract 6.12, Example 4). This is an example of full-class praise, by which the teacher checks in with the students as they work and provides validation. Example 5 (Extract 6.12) shows a similar type of feedback, with the AfL science teacher addressing the work of the class after task completion, noting how much their level was increasing (¨we are going up and up and up!¨). Throughout these examples, a pattern of directed, specific praise is seen with a variety of adjectives used to address the students´ work and achievement. The next extract focuses on examples from the non-AfL classroom.

[13] In this example, the teacher is referring to the students' grasp of the material

Extract 6.13 Non-AfL effective praise

Example 1
Non-AfL TCH: *Well done, everybody, Ok, you, I mean you were able to understand the question very well.*

Example 2
Non-AfL TCH: *Ok, good, STU is making a big effort. Well done.*

Example 3
Non-AfL TCH: *I think you made an effort to think of positive things of wearing a uniform*

The three examples in Extract 6.13 show specific, direct feedback to individual students and the class as a whole. Example 1 (Extract 6.13) shows the teacher praising the students´ responses, referring to their understanding the content. Examples 2 and 3 (Extract 6.13) are directed toward individual students. Instead of praising the students' grasp of the material, the teacher instead gives positive feedback on their effort during the lesson.

6.7.2 Neutral feedback

Neutral feedback encompasses situations in which a teacher offers feedback in an unbiased way.

Neutral feedback Going over the answers of an exercise with the class without communicating any expression of irritation or personal criticism	**AfL frequency** 35	**AfL distribution** 3.85%	**Non-AfL frequency** 51	**Non-AfL distribution** 6.75%	**Chi square** 7.07	**Signif.** +++

The findings reveal a significantly higher distribution of neutral feedback in non-AfL lessons (51/6.75%). Neutral feedback was also found in AfL lessons (35/3.85%). Extract 6.14 shows examples of neutral feedback found in both groups.

Extract 6.14: Neutral feedback in AfL and non-AfL lessons

Example 1
STU*: That, eh, forest is harder than the other.*
AfL TCH: *Yeah, it's hard, isn't it?*

Example 2
STU*: That do aerobics means that you have to raise your hand.*
AfL TCH: *A lot, exactly.*

Example 3
STU: *<L1 History y prehistoria L1>*
Non-AfL TCH*: Yeah, but those are the words in Spanish*

Example 4
STU: *Eh, the prehistory they don't know, they don't know <L1 no saben L1> how to write. I think so*
Non-AfL TCH: *So you think that in prehistory they didn't know how to write, and the difference between prehistory and history is the invention of writing?*

Most examples of neutral feedback found in these classrooms involved the teacher validating or questioning the validity of a student's statement in a neutral way. In many cases, this involved the teacher rephrasing a student's comment (Examples 1 and 4 of Extract 6.14), sometimes adding question tag on the end, or validating the statement. These statements provide a confirmation or rebuttal of the original comment without offering criticism or praise. As seen in Example 3 (Extract 6.14), neutral feedback may also be a way of asking for elaboration, as the student has translated the words "history" and "prehistory" into Spanish when asked to explain their meaning. Thus, neutral feedback is used in a similar way by the AfL and non-AfL teacher, namely as a means of offering validation to students without judgement.

6.7.3 Process feedback

The strategy of giving process feedback allows the teacher to point out mistakes that students have made and help improve future performance by giving advice for improvement.

Process feedback Focusing on what can be learned from the mistakes that have been made, and from the process of producing the correct answer.	**AfL frequency** 6	**AfL distribution** 0.66%	**Non-AfL frequency** 0	**Non-AfL distribution** 0.00%	**Chi square** 5.01	**Signif.** ++

Process feedback represents the smallest percentage of all feedback and one of the most infrequently used strategies in general. This type of feedback was found only in AfL lessons (6/0.66%). Despite the low frequency of usage by AfL teachers, this feedback type plays in an important role in identifying learning gaps, as seen in extract 6.15.

Extract 6.15: AfL Process feedback in science

1 ***STU:*** *The…the sound travels faster than light*
2 ***AFL TCH:*** *You think or you're sure? Because you have to use the book.*
3 ***STU:*** *I'm sure*
4 ***AfL TCH:*** *You can't just give a prediction, you have to base it on facts.*

In the AfL science classroom, a student makes an incorrect hypothesis (Line 1). However, rather than simply correcting the student by providing an answer, the teacher questions whether proper protocol was followed in order to come up with this answer. The teacher had been working with students to train them to become researchers, encouraging them to support all statements with evidence found in their course materials. When the student is asked whether the protocol was followed in this case (Line 2), the response is affirmative, yet the teacher knows that the answer is still incorrect. Therefore, in order to correct the mistake and fill in this knowledge gap, the teacher reminds the class of the protocol for working in science class, which stipulates that predictions must be based on facts (Line 4). Thus, the teacher addresses the learning process rather than the answer, giving the student the tools needed in order to discover this knowledge independently. This AfL strategy that the teacher is following- facilitating inquiry rather than focusing on results- encourages student autonomy.

Process feedback can also be used as a way of pointing out areas for improvement, as seen in extract 6.16 from an AfL citizenship class.

Extract 6.16: AfL Process feedback in citizenship classes

1 ***AfL TCH****: Another thing you can do STU 1 is not just talk about your personal experience*
2 with school, but you can talk in general. School is an institution, all children have to go to
3 school and what school means in general, and then make it more specific. That gives you
4 more chance and gives you more ideas

In this exchange, the students had just completed a speaking activity in which they were given one minute to discuss a topic selected by the teacher: school. This activity was meant to prepare students for a similar speaking task that they would be required to do when taking the Cambridge Preliminary English Test (PET) at the end of the school year. When one student in particular struggles to speak continuously for the full minute, the teacher suggests a way to improve performance rather than pointing out the student's failure to meet the expectations of the required task.

6.7.4 Echo

Echo involves the teacher repeating a student's statement as a way of confirming its validity. Intonation was taken into account to determine whether the repetition was given in a positive or neutral tone. Echo was found frequently in the corpus after preliminary coding and therefore added to the adapted MOLT scheme.

Echo Repeating a statement that the student has made as a means of validation.	**AfL frequency** 104	**AfL distribution** 11.45%	**Non-AfL frequency** 128	**Non-AfL distribution** 16.95%	**Chi Square** 10.39	**Signif.** +++

While echo appeared frequently during all lessons, regardless of the assessment type implemented, the overall percentage was greater in non-AfL lessons (128/16.95%) than AfL (104/11.45%) with a highly significant difference in distribution (chi square= 10.39, $p<0.02$).

Extract 6.17: Examples of echo in AfL and Non-AfL lessons

Example 1
STU: *I don't know*
Non-AfL TCH: *You don't know, ok*

Example 2
STU: *Fossils are living things and the artifacts no*
Non-AfL TCH*: Fossils are living things*

Example 3
STU: *It makes them cry and we have to stop it*
AfL TCH: *We have to stop it*

Example 4
STU: *I don't know how to say it in English*
AfL TCH: *You don't know how to say it*

As observed in Extract 6.17, echo is a means of acknowledging or confirming a statement made by the student and was used in a similar way in AfL and non-AfL lessons.

6.7.5 Summary of feedback

Analysis of the corpus revealed differences in the feedback used in AfL and non-AfL lessons. Effective praise was found more frequently in AfL lessons, with extracts showing specific, direct feedback toward individual students and the class as a whole. Process feedback was found only in AfL lessons, though not frequently. The few instances coded revealed that the use of process feedback helped AfL students identify learning strategies to achieve the set objectives. Neutral feedback was more frequent in non-AfL lessons, mainly with the purpose of validating students' answers without revealing any bias. Echo, which was shown to validate students' statements through repetition by the teacher, was the most frequently found feedback type in both AfL and non-AfL lessons.

6.8 AfL category: Peer and self-assessment

6.8 Peer and self-assessment	6.8.1 Elicitation of peer and self-Correction

6.8.1 Elicitation of self or peer-correction

Peer and self-correction refers to exchanges for the purpose of students evaluating their own or their fellow students' work, which aligns with techniques found in AfL training.

Elicitation of Self-or-peer-correction	**AfL frequency**	**AfL distribution**	**Non-AfL frequency**	**Non-AfL distribution**	**Chi Square**	**Signif.**
Encouraging students to correct Their own mistakes, revise their own work or review/ correct their peer's work	48	5.29%	3	0.40%	33.15	+++

The findings indicate 48 (5.29%) instances of peer and self-correction found in AfL lessons, and 3 (0.40%) found in non-AfL lessons, with a highly significant difference in the distribution (chi square=3.15, $p<0.02$). This motivational strategy was found frequently in the AfL classroom throughout all three academic subjects. To observe the ways in which this motivational technique was used, one extract from each AfL subject is presented, followed by an extract from a non-AfL citizenship class.

As explained in Chapter 2, the use of certain techniques for peer and self-assessment is emphasized in AfL training, especially with young learners. In extract 6.18 the AfL drama teacher engages students in peer-assessment on their peers' drama performance using "two stars and a wish" by commenting on something the students did well and something that must be improved.

Extract 6.18. Two stars and a wish in AfL drama lesson

1 ***AfL TCH:*** *Something to improve on and something they did really well.*
2 ***STU:*** *Me! Me!*
3 ***AfL TCH:*** *Student 1?*

4 ***STU 1:*** *That they did all something at the same time.*

5 ***AfL TCH:*** *So is that something they can improve on or is that something positive?*

6 ***STU 1:*** *They can improve*

7 ***AfL TCH:*** *They could improve on, then. Ok, I'd like a positive comment.*

8 ***STU:*** *That they were funny.*

9 ***AfL TCH:*** *They were funny?*

10 ***STU:*** *Yes*

The teacher asks students to give positive feedback ("stars") and one area for improvement ("wish") to their peers following a theatrical performance (Line 1). One student gives a comment, to which the teacher asks for clarification as to whether this was a positive feature or something to be improved (Lines 4-6). The area of peer and self-correction requires significant training to make students aware of how to deliver feedback appropriately. Therefore, the teacher asks the student to reflect on whether praise or process feedback was being given, which the student clarifies (Line 6) and the exchange moves on to elicit positive feedback. Providing peer assessment serves a duel purpose: the teacher provides a cognitive challenge by encouraging students to reflect upon the criteria while evaluating it, taking on a role of active participation. This technique also motivates students being assessed by giving them positive commentary and areas of improvement for their next performance.

Extract 6.19 shows the same AfL teacher initiating self-assessment with students at the end of a citizenship lesson by referring to initial objectives and reflecting on what they have learned during the lesson. Instead of asking students to discuss their thoughts, the teacher uses the AfL technique of "thumbs up/ thumbs down" to measure their belief of whether the objectives have been reached.

Extract 6.19 Citizenship self-assessment: thumbs up / thumbs down

1 ***AfL TCH:*** *Do you think that looking at the WALT and the WILF, read through those 2 quickly....Student 1, what do you think about today? Do you think we've learned more about 3 how to deal with what happens? Yes, no or so-so? No, you don't. Ok. Student 2, yes? Can I 4 have everybody please? Student 3, you're not showing me how you feel.*
5 *Thank you. This table. So-so, Student 4?*

6 ***STU 4:*** *Yes*

The self-assessment session begins with the AfL teacher displaying the WALT ("We are learning to") and WILF ("What I'm looking for") posters for students, which represent the initial learning objectives. The students are then asked to reflect on whether the objectives have been met by holding their thumb up or down (Lines 2-3). The teacher scans the class to determine which students feel the objectives have been met and those that require further work. The AfL technique thumbs up/ thumbs down facilitates this interaction and avoids generating frustration for students who may have doubts but find difficulty communicating with the teacher. Though the technique does not identify the exact objectives that need further clarification, the teacher now has valuable information for future planning and the students are able to identify gaps in knowledge.

In Extract 6.20, taken from an AfL science lesson, the teacher initiates a peer correction session by exhibiting an individual whiteboard on which a student has written a sentence about the homework and eliciting commentary from the class. This exchange provides an insight into the integration of content and language, as the teacher asks students to comment on the sentences in terms of both.

Extract 6.20: Peer correction session in AfL science lesson

1 ***AfL TCH:*** *Did STU 1 underline the verbs?*

2 ***STU****: All: no.*

3 ***AfL TCH****: No, she didn't, ok, but we're not going to begin attacking, we're going to look at 4 the message, the content, to see if it's correct. Three, two, one, read:*

5 ***STU****: All: Sound is produced by the vibration of an object*

6 ***AfL TCH:*** *Ok, a good sentence about her sentence. Is the content right?*

7 ***STU:*** *All: Yes*

8 ***AfL TCH:*** *Sound is produced by vibrations. So the content is correct. Another positive sentence about her sentence? Student 1?*

9 ***STU 1****: She wrote capital letter and full stop*

10 ***AfL TCH:*** *Ok, capital letter and full stop.*

The exchange begins with the teacher asking the class to comment on the language of the sentence, specifically if the student had followed the instructions for underlining the verbs. Though the class indicates that initial instructions were not followed, the teacher comments that they should not focus on this error, but rather note positive aspects (Lines 2-4). This approach follows the AfL practice of rewarding positive work rather than punishing or focusing on errors. The teacher then assesses the accuracy of content (Lines 6-8) before continuing with a positive comment regarding language (Lines 9-10). Peer correction sessions with individual whiteboards were a common practice in the AfL science classroom, and strengthened the objectives of the unit as well as reinforcing knowledge of grammar.

The final example of peer-and self-correction is taken from a non-AfL citizenship class in which students exchanged notebooks with a partner to correct a dictation.

Extract 6.21. Non-AfL Peer and self-correction: dictation correction

1 ***Non-AfL TCH:*** *We are going to correct the dictation. If you find a mistake, shh, pay*
2 attention please, if you find a mistake, then you underline the mistake. And when we give it
3 back, in pencil, please in pencil, and then look, number one please, you do it with me, check it
4 out.

5 ***STU:*** *(Dictating to teacher): Which city, city, in which city was democracy....*

6 ***Non-AfL TCH:*** *((writing on the blackboard)) Ok, in which city was democracy?*

7 ***STU:*** *Was democracy first established*

8 ***Non-AfL TCH:*** *((writing)): First, check it, please? Shh, ok, if you find a mistake, you*
9 underline it with pencil. Ok?

In this peer and self-correction session, the students exchange notebooks to correct their peers' previously completed dictation. The teacher writes the correct version on the blackboard, instructing students to check their peers' work and underline any mistakes in pencil (Lines 1-4). In this exchange, the instructions focus on identifying mistakes rather than emphasizing positive points or areas of improvement. Thus, though the construct of peer-assessment is followed, the assessment practice follows a summative approach.

6.8.2 Summary of peer and self-assessment

The elicitation of peer and self- correction played a key role in AfL lessons—not surprisingly, as peer and self-assessment represents an important part of AfL practice. Sessions were found throughout all AfL subjects and were usually interaction based, with the teacher eliciting student commentary and providing direction. The approach in most cases was for students to provide positive commentary or areas of improvement, evaluating both content and language driven criteria. The incorporation of AfL techniques was also found in these exchanges, as teachers referenced WALT and WILF as well as "thumbs up/ thumbs down" to reflect on initial lesson objectives and evaluate mastery. Instances of peer and self-correction were not found frequently in non-AfL lessons, and the instance focused on the identification of errors rather than areas of improvement.

6.9 AfL Category: Promoting student engagement

6.9 Promoting student engagement and autonomy	6.9.1. Arousing curiosity or attention 6.9.2. Promoting cooperation 6.9.3 Personalization 6.9.4. Promoting autonomy 6.9.5 Establishing relevance 6.9.6 Scaffolding

6.9.1. Arousing curiosity or attention

Arousing student curiosity or attention is used to generate interest in an upcoming activity.

Arousing curiosity or attention Raising students' expectations that the upcoming activity is going to be interesting	20	2.20%	3	0.40%	9.85	+++

Arousing student curiosity or attention was found more frequently in AfL lessons (20/2.20%) than non-AfL lessons (3/0.40%), with highly significant differences in the distribution (chi square=9.85, $p<0.02$). Many of these examples were brief, yet provoked an intrigued response from students, as seen in extract 6.22 with two examples from science classes.

Extract 6.22: Arousing curiosity

Example 1
AfL TCH*: Imagine we have just entered a cave. This is a cave.*
(students gasp in awe)

Example 2
Non-AfL ***TCH:*** *Imagine that you were living during the Paleolithic period and you were writing in your diary*

One technique found in AfL and non-AfL lessons to arouse attention was asking students to envision themselves in a certain situation to make the material more tangible. In Extract 6.22, Example 1, the AfL science teacher pretends that the class is in a cave when introducing the phenomenon of echo, while in Extract 6.22, Example 2 the non-AfL science teacher asks students to go back in time and imagine that they were living in a pre-historic period. In both cases, the teachers find creative ways of making the material more meaningful to students, provoking an increased interest.

Several examples found in AfL science lessons involve the teacher instructing

students to pretend that they are scientists investigating a certain topic.

Extract 6.23: Arousing attention: good scientists

AfL TCH: *So you have to be good scientists and look for the information that I'm going to write on the blackboard*

The teacher raises expectations for the activity by turning the search for information into a research project, giving students a specific role to play and increasing the meaningfulness of the activity. This AfL practice involves placing students in the role of an active learner with the teacher acting as mediator. Continuing with this role, in the following lesson the teacher introduces the idea of science detectives by leading students in singing the objectives of the next activity to the tune of "Pink Panther", shown in extract 6.24.

Extract 6.24: Arousing attention: science detectives

1 ***AfL TCH:*** *Ready? If you want to laugh, you can laugh, because this is funny.*
2 ***STU****: All: ((turns on music)) science…detectives.*
3 ***AfL TCH:*** *Again!*
4 ***STU:*** *All: Science…detectives ((laugh))*
5 ***AfL TCH****: Ahh*
6 ***STU****: All: Science…detectives*
7 ***AFL TCH:*** *Now the sentence! I'm going to find the information*
8 ***STU****: All: I'm going to find the information, I'm going to find the information, I'm*
9 going to find the information…
10 ***AfL TCH:*** *The next!*
11 ***STU****: All: I'm going to use my brain, I'm going to use my brain, I'm going to use my*
12 brain…

Introducing this musical interlude, which corroborates the overall objectives of the lesson, is designed to enhance interest while introducing expectations; making

students aware of learning aims while arousing their curiosity in the upcoming activity.

6.9.2. Promoting cooperation

Promoting cooperation means engaging students in cooperative work or prompting them to support one another in daily tasks.

Promoting cooperation	AfL frequency	AfL distribution	Non-AfL frequency	Non-AfL distribution	Chi Square	Signif.
Setting up a cooperative learning activity, students encouraging each other	4	0.44%	5	0.66%	0.38	

Instances of activities or comments promoting cooperation made by AfL and non-AfL teachers were not frequent, with 4 instances found in AfL lessons (0.44%) and 5 in non-AfL lessons (0.66%). Opportunities for students to work together were provided in several different ways, though two examples show strategies for cooperation used by teachers. The first occurs in an AfL citizenship lesson in which students are given cards with specific roles to bring a cooperative dimension to a speaking activity.

Extract 6.25. AfL promoting cooperation: assigning roles

> ***AfL TCH:*** *So if it is your turn to present your chart you will have the speaker card. When you are finished, you pass it on to somebody else. Ok, and I'm going to choose an English watcher to make sure you are only talking in English, only English. And one timekeeper, ok? We're going to have… I'm going to give you ten minutes.*

In extract 6.25, students are working in discussion groups as the AfL teacher assigns roles. The members are given a card with a role: the speaker, the English watcher, and the timekeeper. This method, used in cooperative learning environments, provides structure to a group activity and gives students a sense of engagement in the group by requiring active participation. The students are then able to negotiate meaning together in a cooperative way while maintaining a sense of responsibility to their assigned role.

Cooperative activities were also found in the context of group work in the non-AfL citizenship lessons as students engaged in a debate on the pros and cons of wearing a school uniform.

Extract 6.26 Non-AfL promoting cooperation: debate

Non-AfL TCH: *You are going to support, and you are going to defend, one opinion according to the group that you are, so maybe you have to find advantages and disadvantages according to the group.*

Extract 6.26 shows the non-AfL teacher explaining the purpose of the in-class debate in which students are required to defend a stance (pro or con) and work together to list advantages or disadvantages of wearing a uniform. This cooperative learning activity supports the topic of the unit (democracy) by setting up a simulated democratic situation, allowing students to work toward a common goal through a structured debate. Providing an opportunity for students to work cooperatively occurred in AfL and non-AfL classroom environments.

6.9.3 Personalization

Personalization involves providing opportunities for students to express personal experiences to establish a connection to the content.

Personalization	**AfL frequency**	**AfL distribution**	**Non-AfL frequency**	**Non-AfL distribution**	**Chi Square**	**Signif.**
Creates opportunity for students to express personal meanings (feelings, experiences, opinions)	56	6.17%	20	2.65%	11.70	+++

Personalization represents a higher percentage of L2 motivational strategies in the AfL corpus (56/6.17%) than the non-AfL corpus (20/2.65%). The majority of AfL instances were found in the citizenship unit, presumably since it focused on emotions and encouraged students to draw from real life situations. The teacher also allotted many opportunities for students to describe personal reactions to a given emotion, as seen in extract 6.27.

Extract 6.27: AfL Personalization: Emotions

1 ***AfL TCH:*** *And how does it make you feel to be furious?*

2 ***STU 1****: Eh?*

3 ***AfL TCH:*** *How does it make you feel to be furious?*

4 ***STU 1****: Em*

5 ***AfL TCH:*** *What do you feel inside your body?*

6 ***STU 1:*** *Because I am angry*

7 ***AfL TCH:*** *But how do you feel?*

8 ***STU 1****: Very bad.*

9 ***AfL TCH:*** *Bad? Bad how? Who can describe how they feel when they are furious? How do you 10 feel when you are furious?*

12 ***STU 2:*** *A lot of energy*

13 ***AfL TCH:*** *A lot of energy?*

13 (student punches the air, indicating how he would act when furious)

14 ***AFL TCH:*** *Ok, STU 3?*

15 ***STU 3:*** *No*

16 ***AfL TCH:*** *How do you feel when you are furious?*

17 ***STU 3:*** *Eh, I am a lot of... (unintelligible) and I don't talk with anybody.*

In this exchange, the teacher focuses on the emotion "furious", asking students to describe their physical reaction to this feeling. The student does not immediately understand the request, prompting repetition by the teacher, followed by a one-word answer from the student (Lines 1-6). The teacher then poses the question to other students, seeking elaboration and generating more specific responses from others in the class. By giving learners a chance to provide personal meaning beyond a one-word answer, the AfL teacher engages students in an interactive exchange of

guided self-expression. The nature of the unit was particularly conducive to personalization, based on the topic.

The majority of non-AfL instances of personalization were also found in the citizenship unit, as students participated in a debate, working in groups to express personal opinions on the pros and cons of wearing a school uniform, shown in extract 6.28.

Extract 6.28: Non-AfL personalization: wearing a uniform

Non-AfL TCH: *Why should we wear a uniform or why shouldn't we wear a uniform?*
STU 1*: We should wear a uniform because em, we don't waste our time to choose our clothes*
Non-AfL TCH*: Ok, good point.*

The non-AfL teacher asks students to give their opinion on the reasons for and against wearing a uniform. One student responds, giving a positive comment, to which the teacher gives praise (Lines 2-3). By choosing a topic for the debate representing a personal issue for students, the non-AfL teacher is able to generate many opportunities for students to offer their own views. Based on this as well as the previous AfL exchange, it seems that the nature of the lesson as well as the chosen topic play a key role in providing opportunities for personalization.

6.9.4 Promoting autonomy

Promoting autonomy encourages students to take responsibility for their own learning by giving the opportunity to make decisions in class.

Promoting autonomy Offering students choice in activities and making decisions regarding time, internet research	39	4.30%	10	1.32%	12.72	+++

A highly significant difference in the distribution of this subcategory was found in AfL lessons (39/4.30%) when compared to the non-AfL lessons (10/1.32%). This strategy was used in both the regulative and instructional register, suggesting that autonomy was used both as a means of affecting student behavior in the class and

also regarding the content being presented (Christie, 2002). Extract 6.29 gives several examples of promoting autonomy in the regulative register.

Extract 6.29: Promoting autonomy: regulative register

Example 1:
AfL TCH: *I'd like for you to decide who is A and who is B*

Example 2:
AfL TCH: *How many seconds do you have now to draw a thunderstorm?*

****

Example 3:
Non-AfL TCH*: Daniel, I think you are in charge with the cards*

Example 4:
Non-AfL TCH: *You can cut it a bit, feel free to cut the page*

These short excerpts show the ways in which the students are encouraged to be autonomous in daily classroom functions, ranging from choosing partners for pair work to deciding on time limits and being responsible for the materials. Though the contexts differed, the AfL and non-AfL teachers shared the objective of integrating the students into the lesson through assigning roles and encouraging autonomy in making certain choices

Differences occur in the instructional register, in which teachers promote autonomy with a content-based focus. Throughout many of these exchanges, the AfL teachers used referential questions to encourage students to make their own choices regarding their learning.

Extract 6.30: AfL teachers' use of referential questions to promote autonomy

Example 1:
AfL TCH: *Who would like to choose a character?*

Example 2:
AfL TCH: *Do we have to read anything? Or do you think you could just make a prediction just looking at those pictures?*

In these instances in Extract 6.30, the AfL teachers actively involve students by giving them the opportunity to express preferences during the lesson, such as choice of characters to represent or information needed to make their prediction. The use of referential questions allows teachers to provide different avenues of choice.

Providing a forum for making decisions that impact the classroom activities gives students a sense of agency, as their choices determine the flow of the lesson.

Examples of promoting autonomy in content-related situations are also found in non-AfL lessons, though without the use of referential questions.

Extract 6.31: Promoting autonomy in Non-AfL lessons

Example 1:
Non-AfL TCH: *You have to use your imagination*

Example 2:
Non-AfL TCH: *You're going to find the solution*

Example 3:
Non-AFL TCH: *When it's time to vote then we will have to raise the cards saying yes, no or abstain*

In Extract 6.31, Example 1 and 2 show the non-AfL teachers placing students in a position of authority to find the solution to a problem in the classroom rather than simply revealing the answer. In Extract 6.31, Example 3, the non-AfL citizenship teacher gives students the opportunity to vote for or against wearing a school uniform. Through these examples, the students are given authority. However, through referential questions, it could be argued that the AfL teachers are both putting students in this authoritative position and giving them the means of exercising their voice and perhaps influencing the course of the lesson.

Thus, promoting autonomy was found more frequently in AfL lessons, often through the use of referential questions to give students a choice during the activities. The strategy was found in non-AfL lessons as a means of encouraging students to discover the answer to a problem or as a way of participating in the class debate.

6.9.5 Establishing relevance

Establishing relevance occurs when the teacher makes connections between what is being learned and students' everyday lives.

Establishing relevance	AfL frequency	AfL distribution	Non-AfL frequency	Non-AfL distribution	Chi square	Signif.
Connecting what has been learned to everyday lives	23	2.53%	14	1.85%	0.87	

23 instances of establishing relevance were found in AfL lessons (2.53%) and 14 in non-AfL lessons (1.85%), though this difference in distribution is not significant. Instances of establishing relevance were found in AfL citizenship and science lessons and in non-AfL citizenship lessons.

In extract 6.32, the AfL science teacher explains the phenomenon of lightning to students by referencing Harry Potter, a popular culture character with whom most students are familiar.

Extract 6.32: Establishing relevance: Harry Potter

1 ***AfL TCH:*** *Eh, on your whiteboard, please draw the thing on Harry Potter's face ((points to*
2 forehead))

3 ***STU:*** *Ahh!*

4 ***AfL TCH:*** *Ahh! Because I'm pointing here, if you didn't read the books on Harry Potter or you*
5 didn't see the film, maybe you don't know that he has something here. What does he have?
6 ((draws lightning on the board))

7 ***STU****: Scar!*

8 ***STU:*** *Scar*

9 ((all children calling out))

10 ***AfL TCH****: It's one thing…it's one thing*

11 ***STU****: It's a scar*

12 ***AfL TCH:*** *It's a what? It's a scar from what Poww!! ((makes lightning motion))*

13 ***STU:*** *From lightning*

14 ***AfL TCH:*** *From lightning. So on your whiteboard, draw lightning.*

As the teacher introduces the topic of a thunderstorm, she instructs the students to draw lightning on their whiteboards, appealing to the students' knowledge of popular culture by asking them to draw the symbol on Harry Potter's forehead (Lines 1-2). Initially, students are confused by this request, interpreting that they

should draw a scar (Lines 7-11), but through the teacher's scaffolding soon realize that they are being prompted to draw lightning. The teacher uses the example of this popular character to tie the concept of lightning to their daily lives by creating an association anchored in their shared knowledge. This strategy is used to introduce the concept, which is then elaborated upon throughout the unit. By making the initial connection to Harry Potter at the introduction of this phenomenon, the teacher creates a reference point for future lessons.

In extract 6.33, the AfL citizenship teacher connects classroom material to the students' everyday lives through poetry.

Extract 6.33: AfL Establishing relevance: Poem

1 ***AfL TCH:*** *Because the monster's a bit scary and he decides that he's going to check how he*
2 feels, and he says to him ((reading from poem)) 'Are you feeling all right? Have you had a
3 bad day?' has that ever happened to you sometimes when you go home and you've had not a
4 very good day and maybe whoever's at home, if it's your mom, your dad, your brother your
5 grandpa, looks at you and says 'ooh, have you had a bad day?' Has that ever happened to
6 you?

In Extract 6.33, a poem is presented at the beginning of the citizenship unit in which the topic of emotions is explored. The poem describes a scary monster having a bad day. The teacher reads the first few lines and relates the topic to the students' personal experience, inquiring as to whether they have found themselves in the same situation as the monster (Lines 3-6). Through this questioning, the teacher engages the students in the poetry reading and introduces the topic of the unit.

Extract 6.34 gives an example from a non-AfL citizenship class. The students are discussing democracy, which the teacher presents by relating it to the government in their home country.

Extract 6.34: Establishing relevance: Democracy

1 ***Non-AfL TCH:*** *So in our democracy, who is able to vote in Spain? If you sit down properly*
2 you will be able to speak. Good

3 ***STU:*** *((indecipherable))*

4 ***Non-AfL TCH:*** *Children are able to vote?*

5 ***STU:*** *No*

This exchange shows the non-AfL teacher connecting the theme of the unit (democracy) to the system of government in the students' home country. The teacher asks if children are allowed to vote in the Spanish democratic system (Line 4), to which one student replies that they are not (Line 5). Instead of giving students the information, the non-AfL teacher is encouraging them to reflect on how democracy relates to their own lives to produce the answer themselves, which can also be considered an example of effective scaffolding.

6.9.6 Scaffolding

Teachers use scaffolding to facilitate learning by providing support to students in solving a problem or completing an activity.

Scaffolding Providing strategies or models to help students complete an activity successfully (teacher thinks aloud)	**AfL frequency** 77	**AfL distrtibution** 8.48%	**Non-AfL frequency** 43	**Non-AfL distribution** 5.70%	**Chi Square** 4.77	**Signif.** ++

The prevalence of scaffolding was apparent both in the AfL (77/7.48%) and non-AfL classes (43/5.70%), with more examples found in the AfL classroom, generating a trend to significance (chi square=4.77, $p<0.05$). Some examples fall into the category of linguistic scaffolding, in which the purpose was to enable students to correct their own mistakes or facilitate communication. Other examples show the teacher supporting learners to complete an activity, as seen in Extract 6.35 from an AfL citizenship class.

Extract 6.35: AfL: scaffolding in order to complete an activity

1 ***STU:*** *Yes, but I don't know how to do*
2 ***AfL TCH:*** *When you look at that word, what's the first color you think of?*
3 ***STU:*** *Blue*
4 ***AfL TCH:*** *Ok, and thank of one thing that (indecipherable)*
5 ***STU:*** *Because here you know...I don't...*

6 ***AfL TCH****: It's really tricky, but try to think of it*

In this extract, the student comments that they are unsure of how to complete an activity, which requires them to choose an appropriate color for a certain emotion and explain the reasoning behind this choice. The teacher guides the student through the thought process that the activity requires by asking them to take a look at the word and think of the first color that comes to mind (Line 2), and follows up by asking them to explain this choice (Line 4). When the student struggles to communicate the reasoning behind this choice, instead of offering suggestions the teacher encourages the student to try to keep thinking to reach the answer. Scaffolding used for facilitating completion of an activity was found more commonly in AfL lessons, whereas linguistic scaffolding was found frequently in non-AfL lessons. Extract 6.36, taken from a non-AfL citizenship class, is an example of linguistic scaffolding.

Extract 6.36: Linguistic scaffolding

1 ***Non-AfL TCH****: We have one word to just finish the text.*

2 ***STU:*** *One word?*

3 ***Non-AfL TCH:*** *With "F", with "F". Student 1?*

4 ***STU 1****: Finally?*

5 ***TCH:*** *Finally, very good.*

In this instance, the teacher reminds students how to end a composition with the marker “finally”. Since the students are already familiar with it, the teacher offers them the chance to remember the word on their own (Line 1). When the students do not reply (Line 2), the teacher gives the first letter of the word (Lines 3-4). In this way, the teacher is acting as a mediator, encouraging students to be self-reliant in their language skills.

6.9.7 Summary of promoting student engagement

The category of promoting student engagement encompassed several L2

motivational strategies, all of which provided opportunities to increase student attention and involvement during the lessons. The most significant differences when comparing the AfL and non-AfL lessons were found in the categories of arousing curiosity and attention, personalization and promoting autonomy. The AfL science teacher frequently placed students in the role of researchers, which established a sense of interest and autonomy during the lessons. In general, it seemed that AfL teachers gave students the option of making choices regarding the content of the lesson, promoting a sense of autonomy. Personalization was also found with a higher frequency in AfL lessons, though this finding may be attributed to the topic rather than the use of specific strategies. Promoting cooperation and establishing relevance were present in the AfL and non-AfL lessons and used in a similar way by both sets of teachers.

6.10 AfL strategy: Promoting collaboration

6.10 Promoting collaboration	6.10.1 Group Work 6.10.2. Pair Work

6.10.1 Group Work

In group work interactions, students worked with their peers in groups (three or more students) or gave a presentation as a group in front of the class.

Group work The students are mingling, working in fluid pairs, or working in groups (simultaneously or presenting to the whole class).	**AfL frequency** 12	**AfL distribution** 1.32%	**Non-AfL frequency** 5	**Non-AfL distribution** 0.66%	**Chi Square** 1.77	**Signif.**

In AfL lessons, 12 (1.32%) instances of group work were found and in non-AfL lessons 5 (0.66%) instances. While no statistically significant differences were noted, a review of the corpus showed that teachers implemented group work

differently. For reasons of space, table 6.4 gives general descriptions of several types of group interactions found in lessons with examples of each.

Table 6.4 Purpose of group work

Type of Assessment	**Subject**	**Example**	**Purpose**
AfL	Drama	Rehearsing and performing a short play in front of the class	Collaboration
AfL	Citizenship	Discussing emotions. Students were assigned one and had to recall a situation in which they felt this emotion	Discussion
AfL	Citizenship	Retracing chart that students had drawn of good and bad moments over the course of a day and explaining why they had felt that way.	Discussion/ Retroactive Task based
AfL	Science	Teacher gives students one question about homework assignment from previous day and students have to write answers on whiteboards and discuss with classmates.	Discussion/ Retroactive Task based
AfL	Science	Teacher asks students a question and they must be "science detectives" and work in groups to answer using various resources (books, handouts)	Discussion/ Task based
Non-AfL	Citizenship	Teacher gives students an article and each group must summarize one paragraph	Task based
Non-AfL	Citizenship	Teacher gives students questions about an article and they must work in groups in order to answer.	Task based
Non-AfL	Citizenship	In order to prepare for a whole-class debate, students work in groups and prepare pros and cons about wearing a school uniform.	Discussion
Non-AfL	Citizenship	When concluding debate, students discuss in groups whether they will vote in favor of or against wearing a school uniform.	Discussion

Group work played in integral role in the AfL classroom and was present across all subjects. This work almost always involved a communicative activity, for example, discussing the last situation in which students had felt a certain emotion and how

they were able to cope with this feeling in AfL citizenship lessons. At times, these discussions were based on reviewing tasks completed in previous lessons or for homework, elaborating on them or working with classmates to debate a certain aspect facilitated by the teacher. In AfL science lessons, the students engaged in collaborative learning, working as researchers to answer and find evidence for a certain question given by the teacher.

In the non-AfL classroom, group work is found only in citizenship lessons, namely with students working in groups to summarize passages they have read. It was also used during the debate, in which groups of four students were responsible for discussing pros and cons of wearing a school uniform and coming to a consensus to vote.

6.10.2. Pair Work

Pair work refers to situations where students work in groups of two to complete an activity.

Pair work The students are working in fixed pairs (simultaneously or presenting to the whole class).	**AfL frequency** 10	**AfL distribution** 1.10%	**Non-AfL frequency** 2	**Non-AfL distribution** 0.26%	**Chi Square** 4.03	**Signif.** ++

Pair work was found more frequently in AfL lessons, with 10 total sessions (1.10%). 2 instances were found in non-AfL lessons (0.26%), generating a trend in significance (chi square=4.03, $p<0.05$). The purpose of this pair work covers several functions, many of which are similar to those of group work. A selection of these activities is shown in table 6.5, with examples of each type and its purpose.

Table 6.5: Purpose of pair work

Type of Assessment	**Subject**	**Example**	**Purpose**
AfL	Drama	Delivering a monologue about a topic to their partner for one minute to prepare for the PET speaking exam	Discussion
AfL	Science	Looking for definitions in book with a partner	Discussion/ Task based
AfL	Science	Making sounds and describing their characteristics to their partner.	Discussion
AfL	Science	Sharing sentences with one another	Discussion
AfL	Science	Explaining sounds to a partner	Discussion
AfL	Science	Teacher asks a question and students must explain/ discuss answer	Discussion
Non-AfL	Science	Students are given a list of questions related to a Powerpoint that the teacher has shown them and must answer questions in pairs.	Discussion/ Task- based

The majority of examples of pair work were found in science and drama lessons. In the AfL lessons, the science teacher used pair work predominately to encourage students to discuss questions related to the content. Methods for this approach included: presenting students with a concept and opening discussion; evaluating a sentence written previously; or trying to come to a consensus when answering a specific question. This approach was especially useful in encouraging students to discuss an abstract concept, such as 'sound'. Some of these sessions were task based, with the teacher asking students to find a definition or answer questions related to the homework or unit. Instances of pair work found in the non-AfL lesson were also task-based, giving students opportunities to discuss answers to questions related to the content. Finally, in the context of AfL drama, pair work was employed to

facilitate training sessions for the PET speaking exam. These sessions allowed students to practice their speaking abilities on a variety of different topics, mirroring the testing experience.

6.10.3 Summary of promoting cooperation

Group and pair work occurred on a regular basis in both AfL and non-AfL classrooms and served a number of purposes ranging from task-based work to discussions. Based on the results, it appears that group and pair work were found more frequently in AfL lessons, though differences in distribution were present only in pair work. Group work was found in all three AfL subjects, yet only occurred in non-AfL citizenship lessons. The purpose for group work was varied throughout all of the instances found, however the most common examples encompassed task-based work or discussion. In the case of pair work, which was used more frequently in AfL lessons, the purpose was more discussion based and allowed students the opportunity explore content through peer interaction and in some cases prepare for the Cambridge exams.

6.11 Chapter summary

Chapter 6 presented the findings obtained through analysis of the classroom corpus, comparing motivational strategies found in AfL and non-AfL citizenship, science, art and drama units based on frequency, distribution and duration. The findings indicate AfL teachers devoted a higher percentage of class time to L2 motivational strategies in all units. The percentage of time was highest in citizenship units in both AfL and non-AfL lessons. The percentage of strategies used was similar in AfL and non-AfL science lessons. However, the use of strategies by the AfL science teacher was more varied, as opposed to the non-AfL lessons in which fewer strategies were found. Finally, the ratio of strategies was much lower in non-AfL art lessons when compared to AfL drama lessons.

With respect to frequency and distribution, a higher frequency of L2 motivational strategies was found in the AfL lessons. The distribution was also more balanced in the AfL lessons, while the majority of strategies used in non-AfL lessons were from

the categories referential questions and echo. Eliciting peer and self-correction was found frequently in AFL lessons when compared to only three instances in non-AfL lessons and process feedback was found only in the AfL corpus.

This data was supported by an analysis of extracts found for each of the 16 observable motivational strategies in the adapted MOLT framework. This analysis revealed findings such as a more interactive way of signposting used by AfL teachers, facilitated by AfL techniques such as WALT and WILF and "I can" statements. Methods of delivering feedback also differed between the AfL and non-AfL teachers: AfL teachers used more effective praise and integrated some process feedback, while neutral feedback and echo were found more often in non-AfL lessons. Peer and self-correction sessions were found frequently in AfL lessons, requiring active participation of students. These sessions were sometimes facilitated by AfL techniques such as "thumbs up/ thumbs down" and caused students to reflect on learning objectives.

Strategies supporting student engagement were seen more frequently in AfL lessons. In the AfL science lessons, placing the students into the role of active learners with mediation from the teacher led to an increased use of the strategies arousing curiosity and attention and scaffolding. Personalization was seen with a higher frequency in the AfL citizenship lessons, perhaps due to the topic of the unit (emotions). Promoting autonomy occurred more frequently in AfL lessons and gave students a voice in making decisions regarding the content of the unit.

An increased amount of student collaboration in the form of group and pair work was also seen in the AfL lessons, centering on interactive, discussion-based activities such as preparation for the PET speaking exam and discussing the answers to research questions.

Chapters 7 shifts away from classroom discourse to focus on a student-centered approach, measuring the motivation felt by CLIL learners during and after their citizenship lessons and their feelings during specific classroom situations.

Chapter 7: Results: The effects of assessment on student self-reported motivation

7.1. Introduction

The purpose of this chapter is to present data regarding students' self-reported motivation in both AfL and non-AfL citizenship classes. These findings are later compared to the results from Chapter 6 to try to shed light on the relationship between the L2 motivational strategies found and how students described their motivation during the lessons. A small sample of students (N=40 out of 132 participants) was selected to complete the questionnaire at the end of their citizenship unit. The second aim of this part of the study is to determine the feelings of these students in different classroom situations.

The research questions addressed in this chapter are the following:

Research Question 5: Can any relation be seen between the type of assessment used by teachers and student's self-reported motivation? Are students in AfL classes more or less motivated than their non-AfL peers?

Research Question 6: How do AfL and non-AfL students describe their feelings in the context of certain classroom situations?

Participants include one AfL class (n=21) and one non-AfL class (n=19). As described in Chapter 5 (Methodology), the questionnaire was divided into two parts. The first part was an adapted version of the Motivated Strategies for Learning Questionnaire (MSLQ) measuring students' overall motivation and strategies used to maintain this motivation. First, an overview is given of the findings from the five categories of the questionnaire, which include: self-efficacy, test anxiety, intrinsic value, cognitive strategies and self- regulation. Next, a description of the student responses from each item on the questionnaire with an accompanying graph presents the students' responses in more detail.

The second part of the questionnaire used an approach based on an adaptation of a metacognitive template tool in which students evaluated emotions felt in certain

classroom situations (Wall, 2008). These results are presented in graph form with a description for each item.

7.2 Overview of questionnaire results: Parts 1a and 1b

Table 7.1 provides an overview of the results from each category of the questionnaire, comparing the means of AfL and non-AfL student responses. The means were generated by placing a numerical value on each response[14]. The differences between each category were calculated using a two- tailed t-test, to compare the two groups (Brown, 1988). Table 7.1 indicates the mean results of the AfL and the non-AfL group, along with the *p* value.

Table 7.1: Category results of Questionnaire Parts 1a and 1b

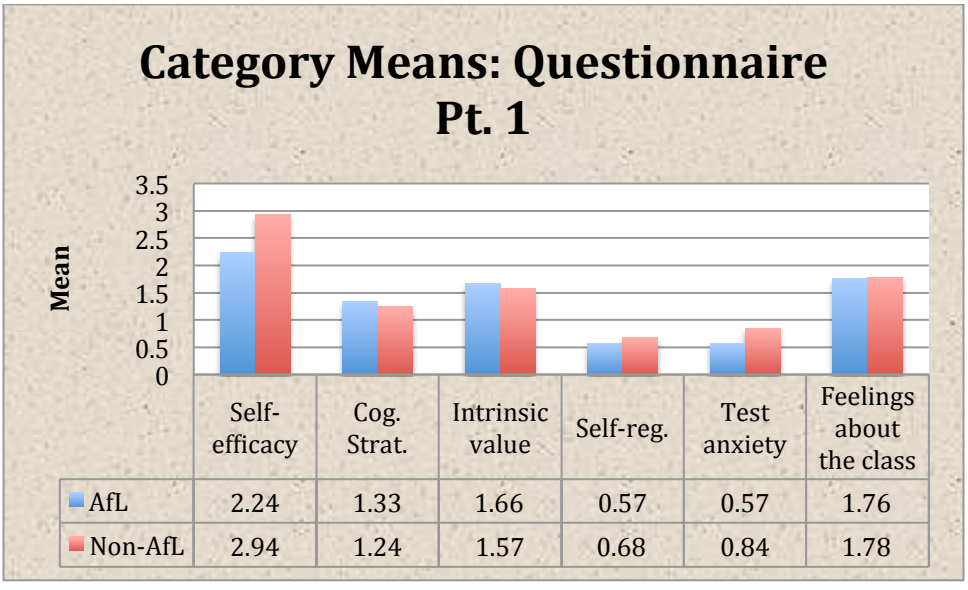
	AfL Mean	**Non-AfL Mean**	***p* value**
Self-efficacy *The extent or strength in belief in one's own ability to complete tasks and reach goals*	2.24	2.94	*0.177 (p>0.05)*
Cognitive Strategies *Mental processes used to accomplish a certain learning goal*	1.33	1.24	*0.551 (p>0.05)*
Intrinsic value *The level of interest of enjoyment students have when engaging in a lesson or activity*	1.66	1.57	*0.683 (p>0.05)*
Self-regulation *Students' abilities to be*	0.57	0.68	*0.474 (p>0.05)*

[14] Numerical values for student responses:
Yes (1)
Sometimes/ I don't know (0)
No (-1)

autonomous or goal driven in their own learning			
Test anxiety REVERSAL *The nervousness a student feels in testing situations*	0.57	0.84	*0.217 (p>0.05)*
Feelings about the class *Measures the students' feelings about engage in classroom activities*	1.76	1.78	*0.839 (p>0.05)*

Figure 7.1 presents the category means for the first part of the student motivational questionnaire (the same data seen in table 7.1) in graph form.

Figure 7.1 Category means: Questionnaire part 1



As seen in table 7.1 and figure 7.1, the means for each category were similar for AfL and non-AfL students. The greatest difference appears in self-efficacy, in which the mean for non-AfL students was higher. The non-AfL student mean for self-

regulation was also higher. The means for AfL students were slightly higher in the case of cognitive strategies and intrinsic value. In test anxiety, the AfL mean was lower, meaning AfL students reported less test anxiety than their non-AfL peers. In the final category, feelings about the class, which referred to the students' attitude toward participating in class and working in groups, the means were nearly the same.

Overall, the results yielded a significance of $p>0.05$ for each category, meaning no significant differences were found between the two groups. However, when considering each individual item on the questionnaire, small differences appear. The next sections present results for individual items in each category of motivation on the first two parts of the questionnaire, comparing AfL and non-AfL student responses.

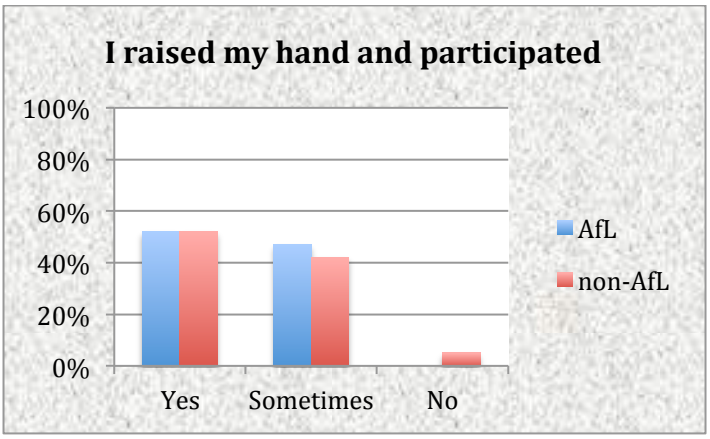
7.3 Student Motivational Questionnaire Results Parts 1a and 1b

The questionnaire results for Parts 1a and 1b encompass students' self-evaluation of their behavior during the class, attitudes toward the course and attitudes toward learning English in general. The categories represented are: self-efficacy, cognitive strategies, intrinsic value, self-regulation, test anxiety and feelings about the class.

7.3.1 Self-efficacy

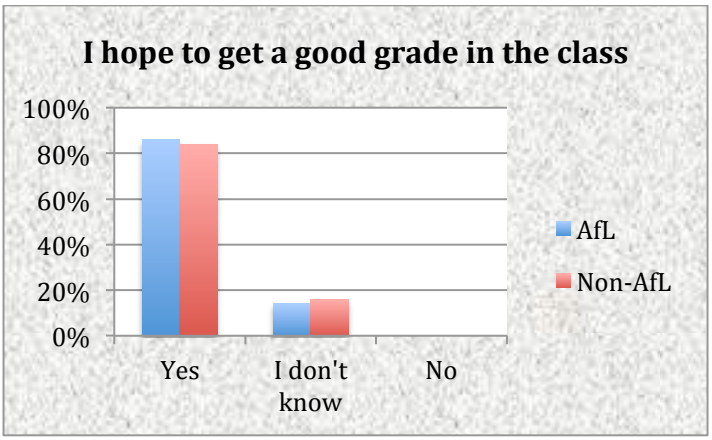
Self-efficacy measures the extent or strength of one's belief in one's ability to complete tasks and reach goals. This category contained the highest number of items on the first part of the questionnaire (n=5).

Figure 7.2: I raised my hand and participated



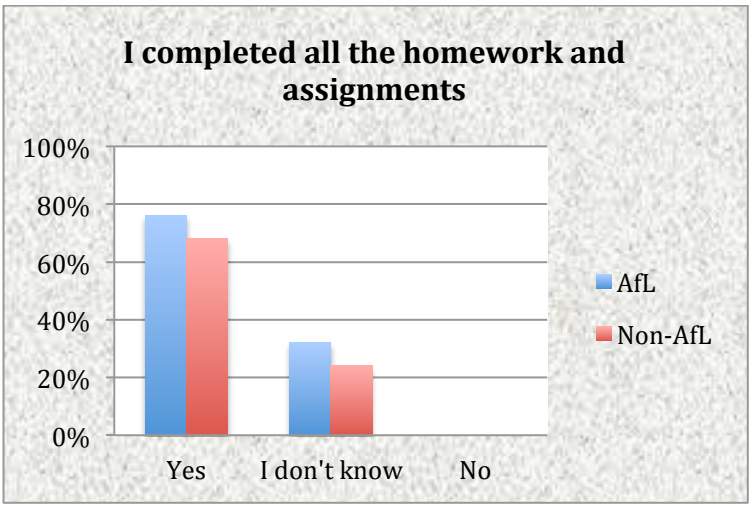
When asked whether they raised their hand and participated in class during the didactic unit, 50% of AfL and non-AfL students responded positively. 47% of AfL and 42% of non-AfL students reported occasional participation. Finally, 0% of AfL students and 5% of non-AfL students responded that they never participated in class.

Figure 7.3: I hope to get a good grade in the class



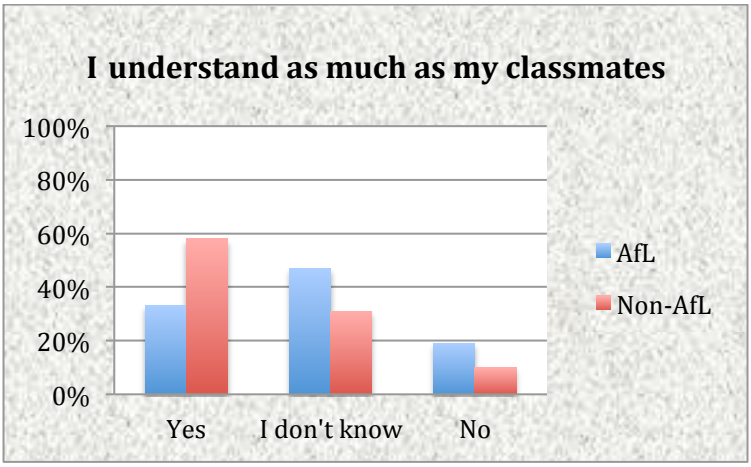
Most participating students hoped to excel in the subject of citizenship, with 86% of AfL students and 84% of non-AfL students reporting expectations for a high grade. 16% of non-AfL students and 14% of AfL students expressed uncertainty regarding whether their grade would be high, while no students in either group reported that they were not hoping for a high grade, as could be expected.

Figure 7.4: I completed all of the homework and assignments



76% of AfL students and 68% of non-AfL students claimed to have competed all homework and assignments from the didactic unit, while 32% of the non-AfL students and 24% of AfL students felt unsure of whether all required tasks had been completed. No students from either class responded that they had not completed their assignments.

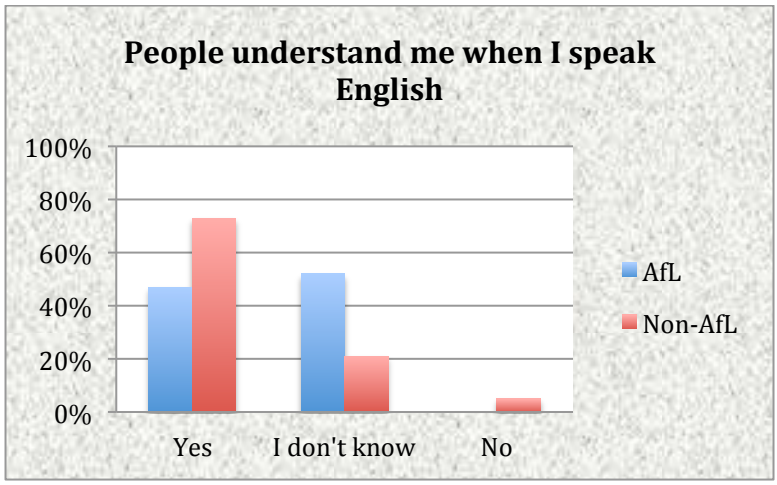
Figure 7.5: I understand as much as my classmates



33% of AfL students and 58% of non-AfL students stated they understood as much as their classmates, while 47% of AfL students and 31% of non-AfL students responded with uncertainty. These results reflected one of the highest negative

responses on the questionnaire, with 19% of AfL students and 10% of non-AfL students admitting to not understanding as well as their peers.

Figure 7.6: People understand me when I speak English



47% of AfL students and 73% of non-AfL students expressed belief that others understood them when they spoke in English. 52% of AfL students and 21% of non-AfL students answered that they were unsure of whether they were understood, and 0 % of AfL students and 5% of non-AfL students responded negatively.

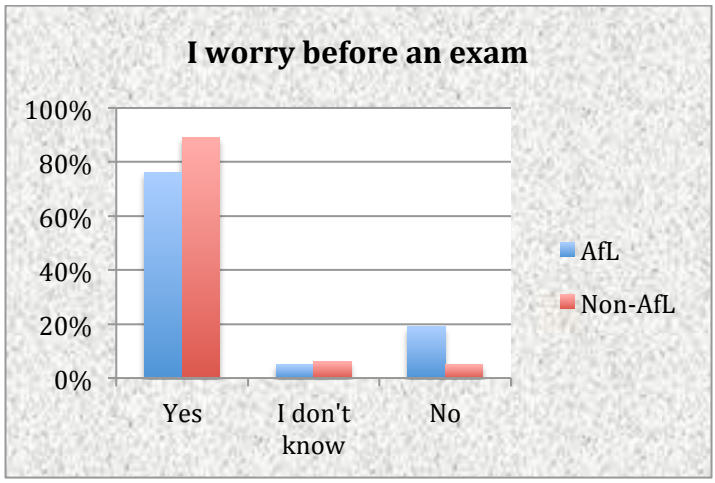
7.3.2 Summary of self-efficacy results

The majority of students from both groups responded positively about the citizenship unit in general when reflecting on their participation and completion of the homework assignments. Most expected to achieve a high grade for the term as a result of this discipline. However, when questioned regarding belief in their own abilities, the non-AfL students reported a higher sense of certainty. The AfL students, on the other hand, expressed uncertainty regarding their English speaking and comprehension abilities when compared to their classmates.

7.3.3 Test Anxiety

Test anxiety measures the amount of nervousness felt in a testing situation and contains only one item. In this case, the question was reversed: students were asked whether they worried before an exam. A positive response in this case is associated with negative feelings.

Figure 7.7: I worry before an exam



76% of AfL students and 89% of non-AfL students did not express worry before the exam, while 6% of the AfL students and 5% of the non-AfL students were uncertain. 19% of AfL students and 5% of non-AfL students reported a lack of worry in an exam situation.

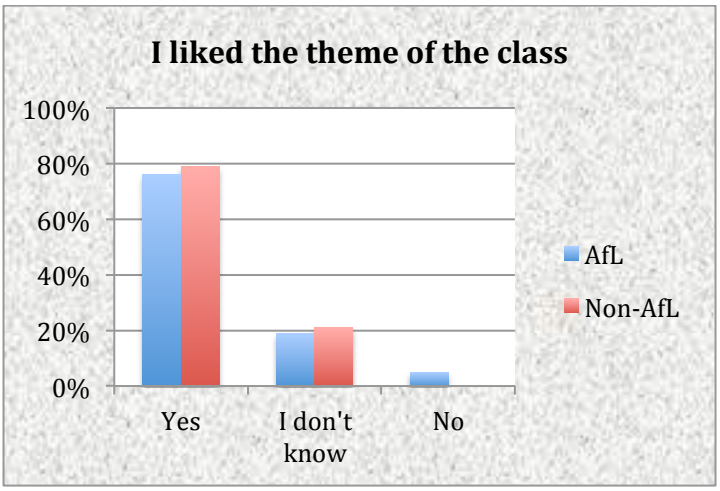
7.3.4 Summary of test anxiety

While both groups of students expressed an increased sense of worry before a testing situation, the percentage was higher in the case of non-AfL students. In the AfL class, more students reported that they felt no sense of worry before taking an exam.

7.3.5 Intrinsic Value

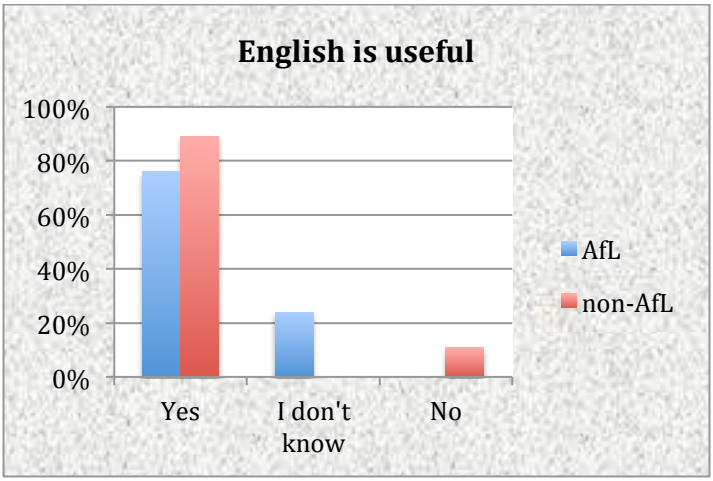
Intrinsic value refers to the level of interest and enjoyment students have when engaging in a lesson or activity. The term may also reflect the perception of whether the completion of this activity is relevant to their lives or will be useful in the future. This category contains two items designed to measure student enjoyment of the theme of the unit as well as their English learning experience.

Figure 7.8: I liked the theme of the class



For the theme of the class, which was 'emotions' for the AfL group and 'democracy' for the non-AfL group, 76% of AfL students and 79% of non-AfL students reported enjoyment. A similar percentage of students were not certain of their enjoyment of the theme (21% of AfL student and 19% of non-AfL students). One AfL student (5%) responded negatively as opposed to 0% in the case of the non-AfL group.

Figure 7.9: English is useful



Most of the students in the non-AfL group believed that English was useful, with 89% of students responding positively and 11% responding negatively. In the case of AfL students, 76% responded that English was useful, while 23% were unsure.

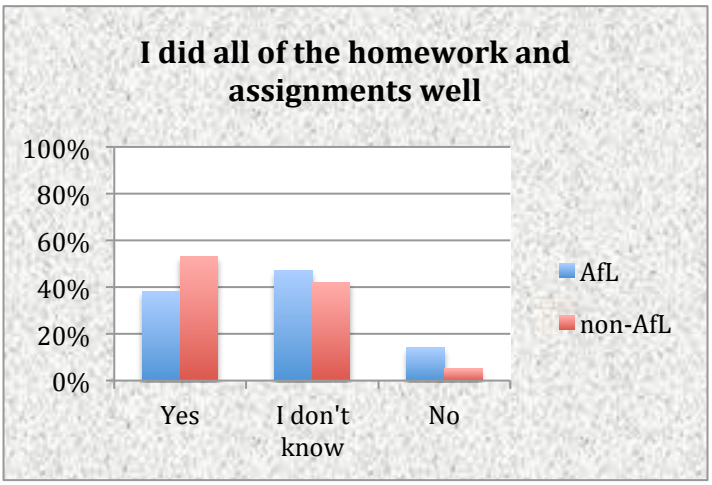
7.3.6 Summary of intrinsic value

The responses of both groups indicate that most students enjoyed the theme of the unit. While the majority of learners indicated that they felt English was useful, a small percentage of the AfL group expressed uncertainty and a few non-AfL students responded negatively.

7.3.7 Cognitive Strategies

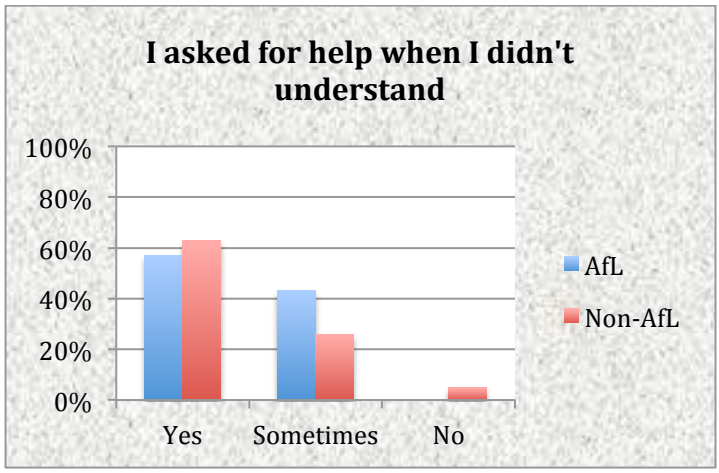
The fourth category focuses on cognitive strategies, or the mental process used to accomplish a specific learning goal. This category contained two items prompting students to reflect on the processes used to carry out cognitive tasks during the citizenship unit.

Figure 7.10: I did all of the homework and assignments well



In this question, students were asked whether they thought they completed the tasks well, to which 38% of AfL students and 53% of non-AfL students provided a positive response. A similar percentage (47% of AfL and 42% non-AfL) responded with uncertainty, and 14% of AfL and 5% of non-AfL students answered that they had not completed the homework well.

Figure 7.11: I asked for help when I didn't understand



57% of AfL students and 63% of non-AfL students reported raising their hand to solicit help from the teacher. 43% of AfL and 26% of non-AfL students reported that they sometimes ask for help when needed, while 0% of AfL students and 5% of non-AfL students reported never soliciting help in class.

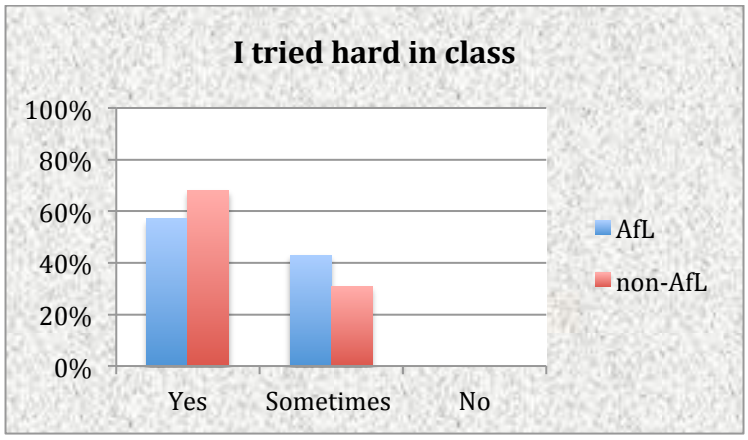
7.3.8 Summary of cognitive strategies

Comparing the student responses for cognitive strategies, the findings are similar for the two groups. Nevertheless, it seems that non-AfL students expressed a higher degree of certainty in their responses. The findings indicate that a higher percentage of AfL students were unsure or did not believe that they had accomplished this goal.

7.3.9 Self-Regulation

The final category on the questionnaire contains only one item: self-regulation. This concept refers to student autonomy and the ability to set and follow their own goals in their learning.

Figure 7.12: I tried hard in class



57% of AfL and 68% of non-AfL students reported putting effort into their performance in class, while 43% of AfL and 31% of non-AfL, admitted that they only sometimes tried their best. No students responded negatively on this item.

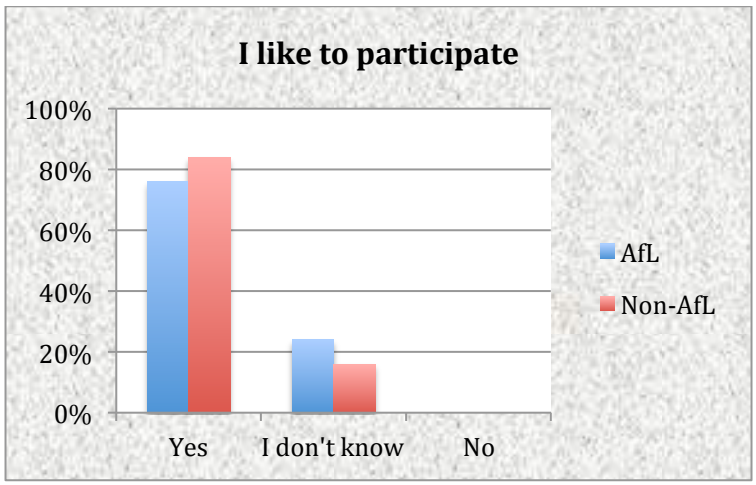
7.3.10 Self-regulation summary

Regarding students' self-evaluation of their ability to regulate their effort during the class, the two groups gave similar responses. These responses indicate that all students believed they worked hard, though this effort may not have been constant.

7.3.11 Students' feelings about the class

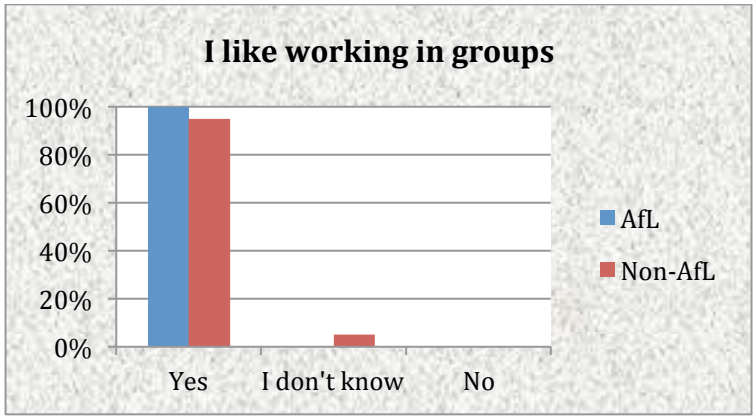
This last category contains items not found on the original MSLQ, which were concerned with evaluating students' feelings about the class in relation to participation and group activities.

Figure 7.13: I like to participate



Most students expressed enjoyment when participating in class, with 76% of AfL and 84% of non-AfL students answering positively. 24% of AfL and 16% of non-AfL students were unsure, and no students answered that they did not like to participate at all in class.

Figure 7.14: I like working in groups



The majority of students in both groups responded that they liked working in groups, with 100% of AfL students and 95% of non-AfL students responding positively. Only 5%, or one non-AfL student, responded with uncertainty regarding their preference for working collaboratively.

7.3.12 Feelings about the class summary

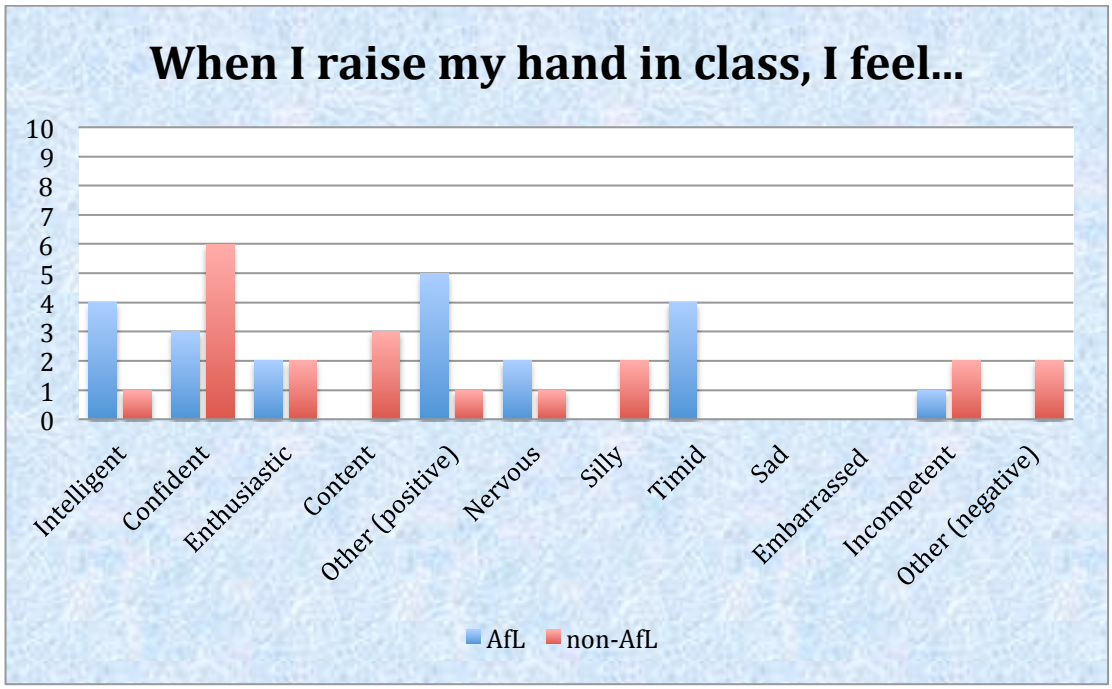
The AfL and non-AfL student responses were similar when reflecting on their feelings about the class regarding participation and working in groups. It appears that both groups of students held a positive view of these classroom activities. This category concludes the first part of the student motivational questionnaire. Section 7.4 provides findings from Part 2 of the questionnaire, in which students reflected on emotions in specific classroom situations.

7.4 Student motivational questionnaire results Part 2

The second part of the questionnaire was designed measure students' feelings based on an adaptation of metacognitive templates in which students were asked to report their thoughts and internal feelings during classroom situations (Wall, 2008). The template was adapted for this study, asking students to choose adjectives to

describe their feelings in certain classroom situations. The results are presented using graphs and include all possible adjectives given, indicating the total number of participants that selected each choice. A section of "other positive" and "other negative" was also included to encompass write-in answers given by some students.

Figure 7.15: When I raise my hand in class I feel…

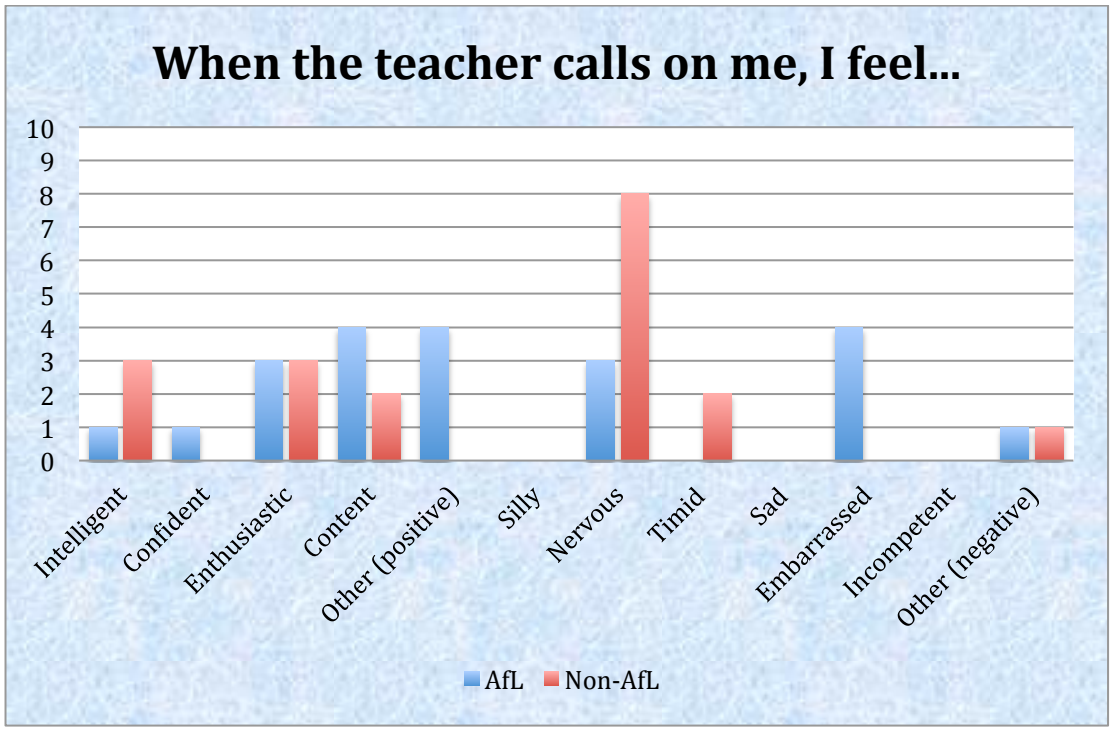


Students responded mainly positively regarding their feelings when raising their hand in class (AfL: +63%/-36%, non-AfL: +66%/-33%[15]).

Some claimed to feel confident (AfL: 3, non-AfL: 6), intelligent (AfL: 4, non-AfL: 1) and enthusiastic (AfL: 2, non-AfL : 2), content (non-AfL: 3) and other positive write-in responses (AfL: 5, non-AfL: 1). However, many students felt unsure about raising their hand, with some selecting nervous (AfL: 2, non-AfL: 1) or silly (non-AfL: 2). Several AfL students (4) reported feeling timid, and others incompetent (AfL: 1, non-AfL 2). There were 2 non-AfL students that wrote in other negative emotions for this situation.

[15] The + reflects the percentage of students who responded positively, the – the percentage of students who responded negatively.

Figure 7.16: When the teacher calls on me, I feel….

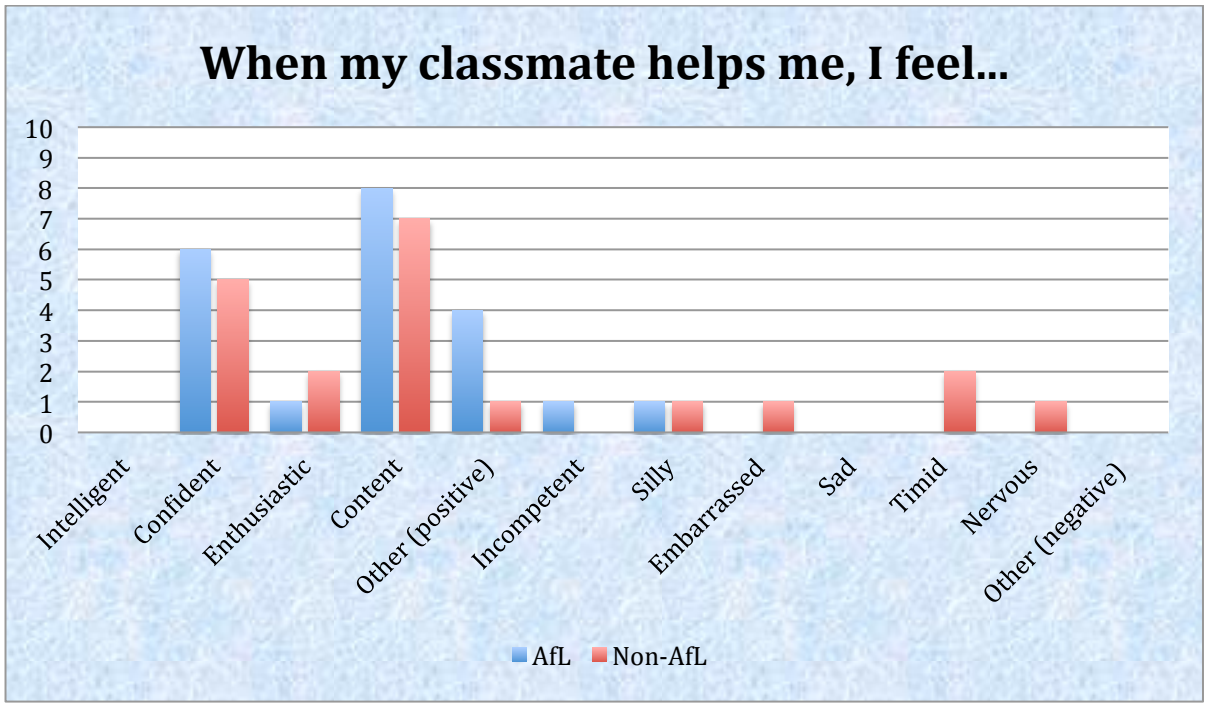


The answers students gave regarding how they felt when the teacher called on them in class were mixed (AfL: +58%/-42%, non-AfL: +48%/-52%[16]).

Some students felt content (AfL: 4, non-AfL: 2) and confident (AfL: 1), enthusiastic (AfL: 3, non-AfL: 3) or intelligent (AfL: 1, non-AfL: 3) and 4 AfL students wrote in other positive adjectives. However, many students felt negatively when the teacher called on them to respond, with the AfL students claiming to be nervous (3), embarrassed (4) and other negative emotions (1). The main emotion that the non-AfL students felt was nervous (8) while some felt timid (2) and 1 wrote in a negative response.

[16] The + reflects the percentage of students who responded positively, the – the percentage of students who responded negatively.

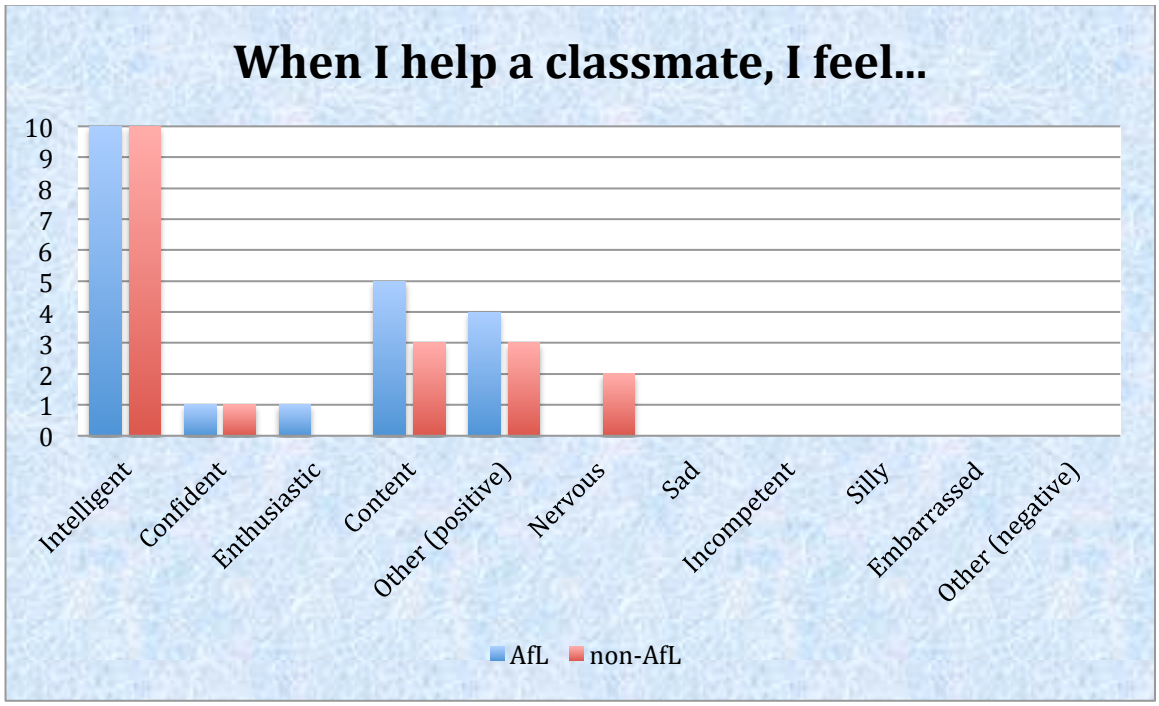
Figure 7.17: When my classmate helps me, I feel…



When students were asked regarding how they felt when a classmate helped them, most gave positive responses (AfL: +90%/ -10%, non-AfL: +77%/-23%).

While no students from either class said they felt intelligent, many expressed confidence (AfL: 6, non-AfL: 5) and contentment (AfL: 8, non-AfL: 7). Some were enthusiastic (AfL: 1, non-AfL: 2) and several wrote in positive adjectives (AfL: 4, non-AfL: 1). While most responses were positive, some students gave negative answers, such as incompetent (AfL: 1), silly (AfL: 1, non-AfL: 1) or embarrassed (non-AfL: 1). Others answered that they felt timid (non-AfL: 2) or nervous (non-AfL: 1).

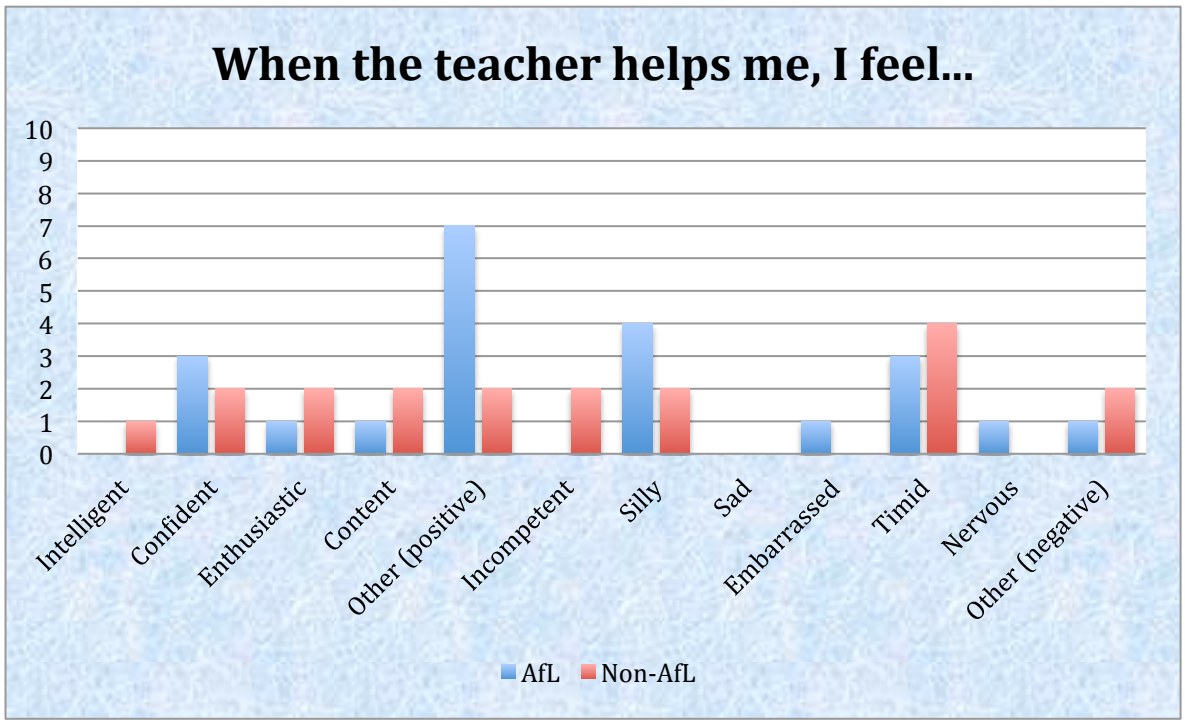
Figure 7.18: When I help a classmate, I feel…



When helping a classmate during the lesson, almost all students expressed positive feelings (AfL: +100%/ -0%, non-AfL: +91%/-9%).

The most common response was intelligence (AfL: 10, non-AfL: 10) or contentment (Afl: 5, non-AfL: 3). Few responded that they felt confident (AfL: 1, non-AfL: 1) or enthusiastic (AfL: 1). Others wrote in other positive adjectives (AfL: 4, non-AfL: 3). The only negative response came from 2 non-AfL students, who expressed feeling nervous when helping their classmates.

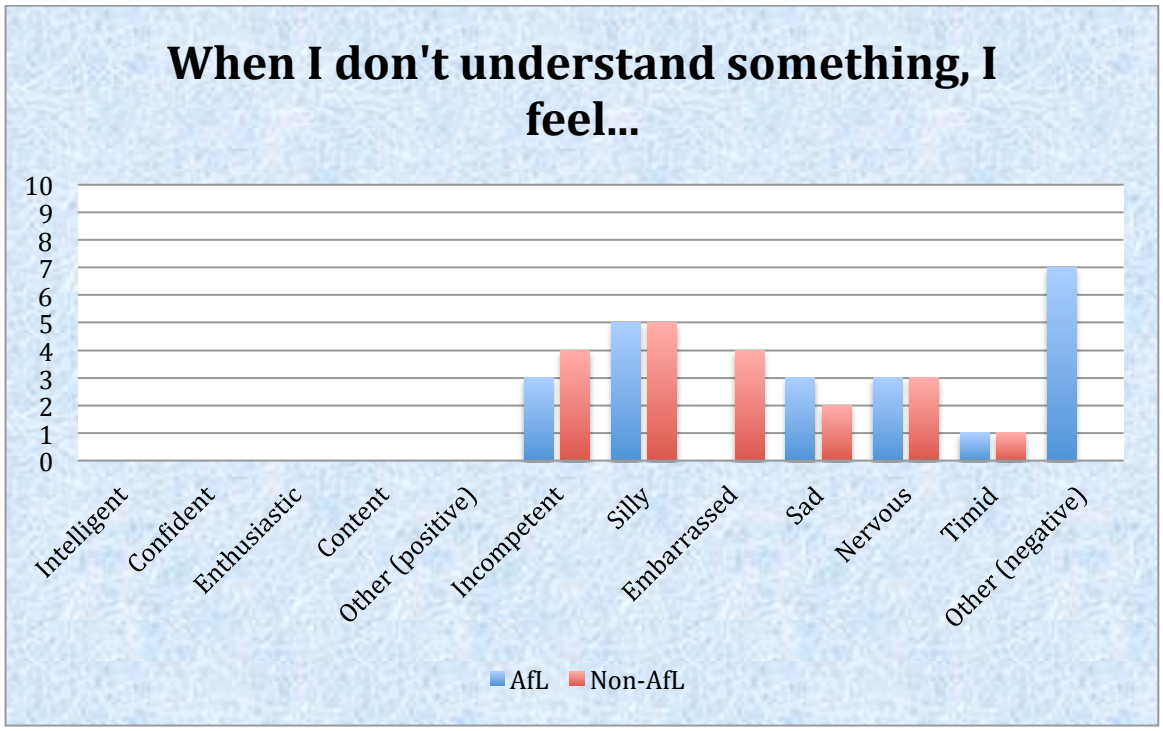
Figure 7.19: When the teacher helps me I feel….



When asked how they felt when a teacher helps them in class, students revealed mixed feelings (AfL: +47%/ -53%, non-AfL: +53/-47%).

Some reported felting intelligent (non-AfL: 1), confident (AfL: 3, non-AfL: 2) and enthusiastic (AfL: 1, non-AfL: 2). Others felt content (AfL: 1, non-AfL: 2) and many expressed other positive feelings (AfL: 7, non-AfL: 2). However, some students responded negatively, saying they felt silly (AfL: 4, non-AfL: 2) or timid (AfL: 3, non-AfL: 4). Some answered that they felt incompetent (non-AfL: 2), embarrassed (AfL: 1) and nervous (AfL: 1) and others wrote in negative responses (AfL: 1, non-AfL: 2).

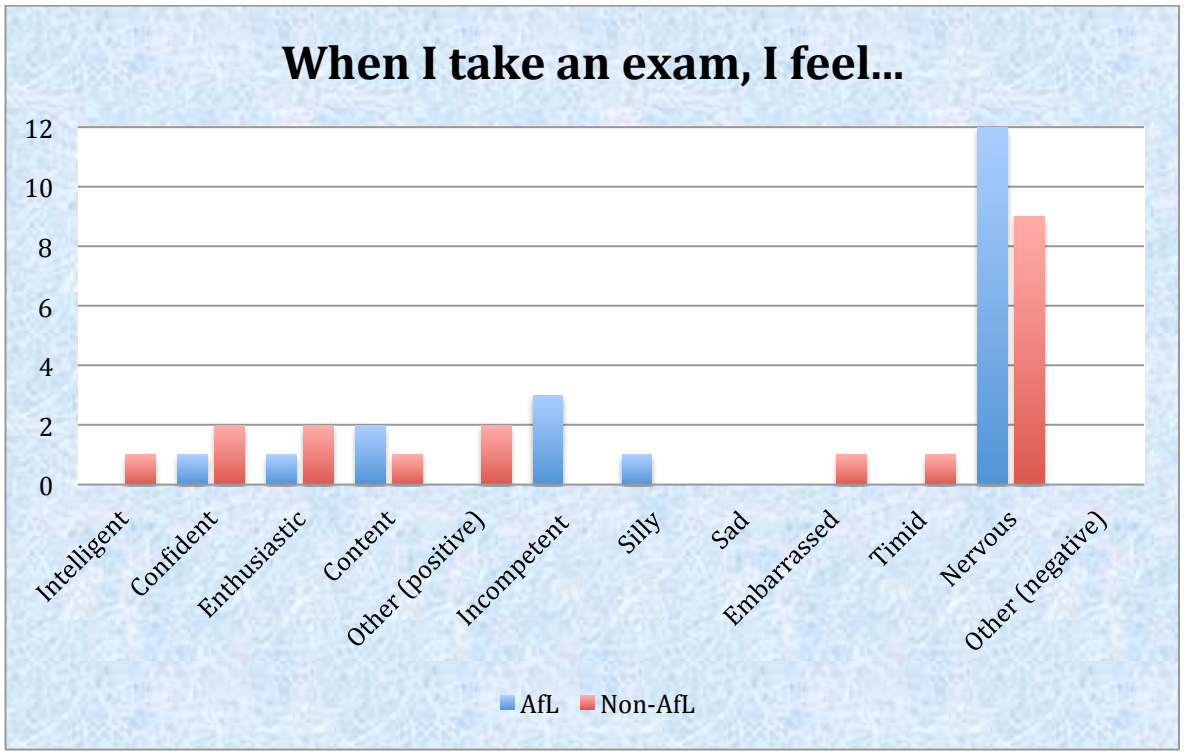
Figure 7.20: When I don't understand something in class I feel…



This item was unique in that it generated entirely negative responses from students (AfL: +0%/ -100%, non-AfL: +0%/-100%).

Understandably, most students do not feel positively when they do not understand something in class. The majority reported feeling silly (AfL: 5, non-AfL: 5) and many also felt incompetent (AfL: 3, non-AfL: 4) or nervous (AfL: 3, non-AfL: 3). This was also the first question in which students answered that they felt sad (AfL: 3, non-AfL: 2) and some also felt embarrassed (non-AfL: 4) and timid (AfL: 1, non-AfL: 1). Many students wrote in other negative adjectives (AfL: 7).

Figure 7.21 When I take an exam I feel…



The final question asked students for their feelings when taking an exam, to which many reported negative feelings (AfL: +37%/ -63%, non-AfL: +48%/-52%).

Overwhelmingly, the majority from both groups claimed to feel nervous (AfL: 12, non-AfL: 9). The remainder of responses varied from student to student, with some choosing intelligent (AfL: 1), confident (AfL: 1, non-AfL: 2), enthusiastic (AfL: 1, non-AfL: 2), content (AfL: 2, non-AfL: 1) and other positive emotions (non-AfL: 2). Others responded negatively, saying that they felt incompetent (AfL: 3), silly (AfL: 1), embarrassed (non-AfL: 1) or timid (non-AfL: 1).

7.4.1 Summary of results: Part 2

Asking students to describe their feelings provides information that might not be visible to a teacher or researcher through classroom observation (Wall, 2008). Reflecting on the responses, it seems that most students felt confident about raising their hand in class, though expressed negative emotions when the teacher calls on them. Almost half of the non-AfL students felt nervous in this situation, and some AfL students felt negatively, as well. When being helped by a classmate, the results

were surprisingly positive despite the fact that this could be thought to reveal weakness, thus a negative reaction from students was expected. The same was true in the case of helping a classmate- most responded positively, though some non-AfL students reported nervousness in this situation. Being helped by the teacher generated a divided response in both groups, with half responding positively and the other half negatively. Not surprisingly, in the case of not understanding, all students responded feeling negatively. The final item, which prompted students to describe how they felt when taking an exam, also generated a negative reaction with few exceptions.

7.5 Chapter Summary

This chapter presented findings on student self-reported motivation through the use of a two-part questionnaire. The first section gave an overview of the category results for the first part (Parts 1a and 1b) with a comparison of the means, which did not show any significant differences between the two groups. A review of each individual category was then provided, elaborating on differences between the AfL and non-AfL citizenship groups for each item on the questionnaire. The chapter continued with a description of the results for Part 2 of the questionnaire, which asked students to reflect on their feelings in certain classroom situations. An interpretation of these findings in relation to those found in Chapter 6 regarding the teachers' use of L2 motivational strategies in the classroom is provided in the discussion chapter (Chapter 9).

Chapter 8: Results: Lower achieving students' reflections on their own learning

8.1. Introduction

This chapter concludes the presentation of the results by analyzing motivational interviews conducted with lower achieving students (N=6) from AfL (n=3) and non-AfL (n=3) citizenship classes. Analysis of these interviews focuses on the learners' evaluation of their own learning process and experience. The interviews were conducted at the end of the citizenship units and encouraged students to comment on their classroom experience in relation to the material, teacher, learning in English in general, and their own perception of their performance. The interviews were analyzed using Martin and White's framework for APPRAISAL (2005), focusing on the two categories of JUDGEMENT and APPRECIATION. The research question for this chapter is the following:

Research Question 7: How do lower achieving AfL and non-AfL students reflect on their own learning and classroom environment?

The chapter begins by providing an overview of the results obtained when analyzing the student interviews. This quantitative data is supported with qualitative data analyzing examples of extracts in which students use APPRAISAL to reflect on their learning. The extracts reflect the annotation method, which took into account units longer than a single token and instead annotated full sentences.

8.2 Overview of judgement and appreciation

The interviews were conducted using an adaptation of the Student Motivational State Questionnaire. [17] Table 8.1 shows the frequency of APPRAISAL found in the six interviews [For full interview transcript see Appendix 9]. The findings indicate that the students used both JUDGEMENT and APPRECIATION throughout the interviews.

[17] The interviews lasted approximately 5-7minutes with each student, with the word count for the transcriptions totaling 6.452 words.

Table 8.1: Appraisal results

	AfL		non-AfL			
Feature	N	Percent	N	Percent	Chisqu	Signif.
APPRAISAL	N=61		N=68			
attitude	61	100.00%	68	100.00%	0.00	
ATTITUDE	N=61		N=68			
judgement	29	47.54%	33	48.53%	0.01	
appreciation	32	52.46%	34	50.00%	0.08	
EXPLICITNESS	N=61		N=68			
inscribed	30	49.18%	40	58.82%	1.205	
invoked	31	50.82%	28	41.18%	1.205	

APPRAISAL was found throughout the student interviews with 61 instances among AfL students and 68 among non-AfL students. There were no significant differences found between groups. This was perhaps due to the effects of the instrument, which encouraged students to share their views in a structured way, with the nature of the interview questions prompting a degree of APPRAISAL. A similar frequency of JUDGEMENT (AfL: 29/ non-AfL 33) and APPRECIATION (AfL: 32/ non-AfL: 34) was found in each group. The same was true for the instances of INSCRIBED (AfL: 30/ non-AfL: 40), or explicit, and INVOKED (AfL: 31/ non-AfL: 28) or implicit APPRAISAL. The next sections elaborate on the results, showing how JUDGEMENT and APPRECIATION were used and what they revealed about the learning processes undergone by each group of students.

8.3. Judgement

Judgement	Evaluates human behavior ethically (morally and legally). Language which criticizes or praises, condemns or applauds the behavior - actions, deeds, sayings, motivations, etc. – of human individuals or groups.

The six sub- categories of JUDGEMENT are: NORMALITY, CAPACITY, TENACITY, PROPRIETY, VERACITY and UNCLEAR (results that did not fit into one of the five categories defined by Martin and White). Results are shown in table 8.2.

Table 8.2 Overall results: JUDGEMENT

JUDGEMENT	AfL		Non-AfL		ChiSq	Sig.
normality	0	0.00%	0	0.00%	0.00	
capacity	19	31.15%	24	35.29%	0.25	
tenacity	4	6.56%	3	4.41%	0.29	
propriety	1	1.64%	0	0.00%	1.12	
veracity	0	0.00%	0	0.00%	0.00	
unclear	5	8.20%	6	8.82%	0.02	

The majority of examples of JUDGEMENT were found in the category of CAPACITY in both the AfL (19/31.15%) and the non-AfL interviews (24/ 35.29%). TENACITY was also used, though less frequently, with 4 (6.56%) instances in the AfL interviews and 3 (4.41%) in the non-AfL interviews. Only one (1/1.64%) instance of PROPRIETY was found in the AfL group. No instances of NORMALITY or VERACITY were found in the corpus. The remaining examples (AfL: 5/8.20%, non-AfL: 6/8.82%) were classified as UNCLEAR after discussion with a UAM professor dedicated to researching APPRAISAL. Sections 8.3.1-8.3.6 describe the positive, negative and neutral instances in each sub-category, with accompanying extracts to show how students used APPRAISAL to reflect on their learning.

8.3.1 Social esteem: Capacity

CAPACITY, or making an assessment of another or one's own ability or skill, represented the highest frequency of JUDGEMENT made by students. Each instance was further classified into whether the students expressed CAPACITY in a positive, negative or neutral way (see table 8.3).

Table 8.3: Capacity

CAPACITY	AfL		Non-AfL		ChiSq	Sig
positive-capacity	10	16.39%	12	17.65%	0.04	
negative-capacity	8	13.11%	11	16.18%	0.24	
neutral-capacity	1	1.64%	1	1.47%	0.01	

Students expressed a similar amount of positive CAPACITY, with non-AfL students expressing slightly more than AfL students (AfL: 10/16.39%, non-AfL: 12/17.65%). The negative examples were also higher in the case of non-AfL students (AFL: 8/ 13.11%, non-AfL: 12/17.65%) with neutral example from each group (AfL: 1/1.64%/ non-AfL: 1/1.47%). The comments in which AfL students expressed the highest level of CAPACITY were regarding their ability to speak Spanish in contrast to English, as seen in extract 8.1.

Extract 8.1: AfL: Positive capacity and the ability to speak Spanish

Q: *Do you feel more nervous in classes taught in English or Spanish?*

AfL STU 2: *I already know how to speak Spanish well*

AfL STU 3*: And in Spanish language, for example, if I have to do a story in Spanish I don't worry a lot because I know how to read well*

In extract 8.1, the AfL students exhibited a high amount of confidence in their abilities to express themselves and read in the Spanish language, which was to be expected considering that all were native Spanish speakers.

On the other hand, AfL students expressed doubts in their English abilities. Extract 8.2 shows AfL negative CAPACITY. The extract is divided into three examples based on students' responses to three different interview questions.

Extract 8.2: AfL negative capacity and the ability to speak English

Example 1
Q: *Do you think that one day you will be capable of speaking English very well?*
AfL STU 2*: I don't know, maybe if I study a lot of English, yes*

Example 2
Q: *Do you worry a lot about making mistakes when you are speaking English?*
AfL STU 1: *Yes, because some things, maybe, eh, well, saying instead of "team", meaning team, well, instead I say /taim/ like the time, and, and....that, like the...and sometimes I make mistakes like (the words) are almost same like "mouse" and "house" and instead of saying "mouse" I say "house"*

Example 3
Q: *In the classes that you have in English, normally do you understand what you need to do and how you need to do it?*
AfL STU 2*: Eh, well, sometimes I don't. Because she says some things and eh, I don't understand them.*

In Example 1 of Extract 8.2, students were asked to comment on their future capacities to speak English. One AfL student remains ambiguous about being capable of speaking English well in the future, and qualifies this statement by saying that a great deal of study is required in order to achieve this goal. In Examples 2 and 3, the students express specific insecurities regarding mistakes made in English and their comprehension during the lesson. The non-AfL students express similar doubts in their own English speaking abilities, as seen in Extract 8.3.

Extract 8.3: Non-AfL negative CAPACITY: making mistakes in English

Example 1
Q: *Do you worry a lot about making mistakes when you speak English?*
Non-AfL STU 1*: I don't know…I worry…I don't know*
Non-AfL STU 3: *Yes because I'm not very confident and I worry that I don't know English very well.*

Example 2
Q: *Are you scared that your classmates will laugh at you when you are speaking English in class?*
Non-AfL STU 2: *Sometimes. Because I'm embarrassed everyone will laugh at me.*

The non-AfL students talk openly about their perceived lack of capacity for speaking English, expressing insecurity and worry about their speaking abilities. In Example 2 of Extract 8.3, one non-AFL student indicates a sense of embarrassment.

However, most non-AfL students responded positively when asked to predict their grades in their English lessons, indicating that they would excel.

Extract 8.4: Non-AfL positive capacity

Q: *Do you think that you're going to get good grades in your classes taught in English?*
Non-AfL STU 1: *Yes*
Non-AfL STU 2: *Yes*

In extract 8.4, non-AfL students anticipate achieving good grades in the classes taught in English, which contrasts insecurities and worry seen in extract 8.3. On the other hand, the AfL students appear doubtful of their potential to pass their classes and achieve high grades, citing specific difficulties such as behavior, trouble with certain grammar points and comprehension difficulties.

8.3.2. Social esteem: Tenacity

TENACITY refers to a person's determination or resolve. In this case, TENACITY relates specifically to a students' ability to be hard working or make an effort in the classroom.

Table 8.4: Tenacity

TENACITY-TYPE	AfL		Non-AfL		ChiSq
Positive-tenacity	3	4.92%	2	2.94%	0.34
Negative-tenacity	1	1.64%	1	1.47%	0.01

Few examples of TENACITY were found in the interviews, with some positive examples (AfL: 3/ 4.92%, non-AfL: 2: 2.94%) and only one negative response for each group (AfL: 1/1.64%, non-AfL: 1: 1.47%). The majority of these instances were given when students were asked if they worked hard in their English classes to please their teacher.

Extract 8.5: Tenacity

Q: *Do you work hard in English class to make the teacher happy?*
AfL STU 1: *Yes*
AfL STU 2*: Yes*
AfL STU 3: *Yes*

Non-AfL STU 1*: Enough*
Non-AfL STU 2*: Well, sometimes yes and other times no. I don't know, because sometimes I get a little bored.*
Non-AfL STU 3*: Yes*

As seen in extract 8.5, while the majority of responses given were positive, one non-AfL student showed ambivalence in describing work ethic, citing boredom as one of the main reasons.

8.3.3. Social sanction: Propriety

PROPRIETY represents an evaluation of ethical or moral behavior of a person. During the interview, students did not make use of this type of evaluation very often: there was only one instance of negative PROPRIETY.

Table 8.5: Propriety

PROPRIETY	AfL		Non-AfL		ChiSq
positive-propriety	0	0.00%	0	0.00%	0.00
negative-propriety	1	1.64%	0	0.00%	1.12

This comment, made by an AfL student, was in regards to whether or not they anticipated passing subjects taught in English.

Extract 8.6: Propriety

Q: *Do you think that you are going to pass the subjects taught in English*
AfL STU 3: *Yes, I think so. I don't know, because there are some classes that I don't behave really well in.*

The student first responds positively to the question, but then reflects on all elements that might lead to a passing grade, such as behavior. While the response might reveal insecurity, it is also indicative of the student's ability to consider all criteria that might be taken into account when calculating final grades (i.e. behavior, in addition to homework and exams). The fact that the student was honest about a perceived weakness reflects training in targeting areas of improvement.

8.3.4. Unclear

The remaining examples of JUDGEMENT were ambiguous, making it difficult to group them into one of the sub-categories. Instances classified as UNCLEAR was due to the fact that they were indirectly related to another category and shown in table 8.6.

Table 8.6: Unclear judgement

UNCLEAR	AfL		Non-AfL		ChiSq
positive-unclear	3	4.92%	5	7.35%	0.33
negative-unclear	2	3.28%	1	1.47%	0.46

Extract 8.7: UNCLEAR instances of judgement

Q: *Do you feel more nervous in classes taught in English than classes taught in Spanish?*

AfL STU 3: *Yes because…because more nervous when the teacher calls on me to come or tell a story or something, there I get tongue-tied.*

In extract 8.7, an AfL student indicates a sense of nervousness felt in English classes, resulting in difficulty in expression. While this declaration is perhaps indirectly related to CAPACITY, the catalyst of nervousness suggests the experience is situational rather than tied to the student's abilities.

8.3.5 Categories of Judgement not found in student interviews

Instances of the categories NORMALITY, or the assessment of how special or unusual a person's behavior is and VERACITY, an evaluation of a person's honesty or moral value, were not found in the student motivational interviews.

8.3.6. Summary of judgement

Analysis of JUDGEMENT in the lower achieving student motivational interviews showed instances in the categories of TENACITY and CAPACITY, which included self-evaluations, mainly related to capabilities and drive for English language learning. When reflecting on CAPACITY, the category of JUDGEMENT found most frequently,

students gave both positive and negative self-assessments. No instances of NORMALITY or VERACITY were found, and only one instance of PROPRIETY was coded. Some instances of JUDGEMENT found which did not fit into any specific category in the Martin and White framework were coded as UNCLEAR. These instances were found to be situational (i.e. feeling nervous in class) rather than relating directly to intrinsic qualities.

The next section describes instances of APPRECIATION, or the evaluation of process, performances and other elements unrelated to human behavior.

8.4 Appreciation

Appreciation	Evaluates things, processes and states of affairs aesthetically or with the social value accorded to the object. Human participants may be 'appreciated' where the assessment does not directly focus on the correctness or incorrectness of behavior, but rather the appearance of the person

Table 8.7: Appreciation results

APPRECIATION-TYPE	AfL		Non-AfL		ChiSq
reaction	11	18.03%	9	13.24%	0.56
composition	11	18.03%	9	13.24%	0.56
social valuation	10	16.39%	16	23.53%	1.02

Half of the instances of APPRAISAL found in these interviews were coded as APPRECIATION (AfL N=32, non-AfL N=34). APPRECIATION was found in instances of students evaluating their course content, their experience learning in English and the language learning process in general. The total amount of APPRECIATION coded was very similar in AfL and non-AfL interviews: REACTION (AfL: 11/18.03%, non-AfL: 9/13.24%), COMPOSITION (AfL (11/18.03%, non-AfL: 9/13.24%) and VALUATION (AfL: 10/16.39%, non-AfL: 16/26.56%).

8.4.1. Reaction

REACTION refers to an emotional response describing an event itself rather than the feeling of the speaker. These descriptions can be classified as either IMPACT (How does it strike me?) or QUALITY (Do I like it?) (see table 8.8).

Table 8.8: Reaction

REACTION	AfL		Non-AfL		ChiSq
impact	3	4.92%	4	5.88%	0.06
quality	8	13.11%	5	7.35%	1.18

Examples of IMPACT were slightly higher in the non-AfL interviews (AfL: 3/4.92%, Non-AfL: 4/5.88%) while examples of quality were found more in the AfL interviews (AfL: 8/13.11%, Non-AfL: 5/7.35%).

8.4.1.a. Impact

IMPACT is used to describe the initial reaction a situation or experience might have on the experiencer. Instances of IMPACT articulated by students were classified further as positive, negative or neutral, as seen in Table 8.9.

Table 8.9: Impact

IMPACT	AfL		Non-AfL		ChiSq
positive-impact	2	3.28%	3	4.41%	0.11
negative-impact	1	1.64%	1	1.47%	0.01
neutral-impact	0	0.00%	0	0.00%	0.00

Very few instances of IMPACT were found in the students' responses during the motivational interviews. Of these examples, some were positive (AfL: 2/3.28%, Non-AfL: 1/4.41%) and others negative (AfL: 1/1.64%, Non-AfL: 1/ 1.46%) with no neutral instances. The AfL students mainly focused on the IMPACT citizenship classes had on them, expressing their reaction in both a positive and negative way.

In extract 8.8, AfL students are asked about their reactions to being in citizenship classes and whether they would choose for the classes to continue a bit longer.

Extract 8.8: AfL IMPACT of citizenship classes

Q: *When your citizenship classes end, do you sometimes wish they would go on a bit longer?*
AfL STU 1: *There are times when it's a little bit boring.*
AFL STU 3: *Yes. Because there are some things that are very entertaining*

The first student response is negative, citing certain moments of the class which are perceived as boring, while the second student finds the opposite to be true. It is important to note the explicitness of these declarations, though the reasoning for the evaluation is not stated.

Examples of IMPACT by non-AfL students were expressed mainly in relation to learning in Spanish versus English.

Extract 8.9: Impact: learning in Spanish versus English

Q: *Do you prefer learning in English or in Spanish?*
Non-AfL STU 1*: I prefer Spanish, but in English there are better things*

This AfL student initially expresses a preference for Spanish followed by a positive evaluation of English in general. While the student does not elaborate on the "better things," the commentary suggests a positive valuation of class activities in English.

8.4.1.b. Quality

QUALITY refers to an emotional reaction or response to a particular event or situation. The majority of these instances found describe students' feelings toward learning in English (see table 8.10).

Table 8.10: Quality

QUALITY	AfL		Non-AfL		ChiSq	
positive-quality	6	9.8%	4	5.88%	0.70	
negative-quality	0	0.0%	0	0.00%	0.00	

neutral-quality	2	3.2%	1	1.47%	0.46	

Table 8.10 shows that expressions of QUALITY were mostly positive (AfL: 6/ Non-AfL: 4) with no negative sentiments and few neutral examples (AfL: 2/ Non-AfL: 1). The evaluations are mostly concerned with students' attitudes toward learning in English.

Extract 8.10: AfL: Quality in relation to learning in English

Q: *Do you like to learn in English?*
AfL STU 1: *Yes because…um, the….because the words are different and….and I learn new things*
AfL STU 2: *Yes, to…because this way I can learn more*

Non-AfL STU 3: *Yes. Well, I like English a lot but it's really hard for me to learn it, but I like it.*

In general, the students expressed enjoyment for learning in English, with three students responding affirmatively while qualifying their answers. The first AfL student refers to the increasing vocabulary and new knowledge more generally. The second AfL student cites a higher amount of learning, though it is unclear whether this refers language or content. Finally, the non-AfL student also gives a positive APPRECIATION of learning in English, showing a positive attitude toward English, though qualifying the statement by recognizing the difficulty involved. The responses made by the AfL students suggest an understanding of the language element in the CLIL classroom, which leads to an increase in learning through the addition of the foreign language. This is also mentioned by the non-AfL learner, whose opinion of English remains positive despite the added complexity of learning in a bilingual context.

8.4.2. Composition

COMPOSITION refers to an assessment of the evaluator's perception of something based on viewing the parts as a whole and how they fit together. Of the two sub-categories, only COMPLEXITY was found.

Table 8.11: Composition

COMPOSITION	AfL		Non-AfL		ChiSq
balance	0	0.00%	0	0.00%	0.00
complexity	11	18.03%	9	13.24%	0.56

Several instances of COMPLEXITY were used by the AfL (11/18.03%) and non-AfL students (9/13.24%). The main purpose for students commenting on COMPLEXITY referred to English or learning in English, with some references to the content of their citizenship classes.

8.4.2.a Complexity

COMPLEXITY in this case refers to the students' stance on how intricate or difficult a certain concept is to follow. This sub-category is reflected in positive (simple) neutral (not simple nor complex) and negative examples (very complex) indicated by the interviewees, as shown in table 8.12.

Table 8.12: Complexity

COMPLEXITY-TYPE	AfL		Non-AfL		ChiSq
positive-complexity	3	4.92%	2	2.94%	0.34
negative-complexity	7	11.48%	3	4.41%	2.24
neutral-complexity	1	1.64%	4	5.88%	1.55

While not many examples of COMPLEXITY were found, the examples in the case of AfL students reflect negative COMPLEXITY (7/11.48%). Some references to positive COMPLEXITY (3/5.92%) and one reference to neutral COMPLEXITY (1/1.64%) were found in AfL interviews. Non-AfL students used some examples of positive (2/2.94) and negative COMPLEXITY (3/4.41%), and a higher frequency of neutral COMPLEXITY (4/5.88%). The majority of these instances refer to the challenges of learning in English, mainly expressed by AfL students.

Extract 8.11: AfL: Difficulties of learning in English

Example 1:
Q: *Do you think that learning in English is easy or difficult?*
AfL STU 1: *Yes. Well, it's difficult because I don't know much English well yet. It's a little hard for me.*
AfL STU 2: *Um, at the beginning. But when you don't know anything, well, it appears difficult to you.*
AfL STU 3*: Mmm, more than in Spanish*

Example 2:
Q: *Do you think that one day you will be capable of speaking English very well?*
AfL STU 1*: I don't know because it's very difficult.*

Example 3:
Q: *Do you feel more nervous in classes taught in English than classes taught in Spanish?*
AfL STU 1: *(sometimes yet and others no) because some are so....a little difficult and...and...and in math for example I understand better*

In Extract 8.11, Example 1, the first student responds that learning content in a foreign language is difficult due to having a lower English level. The second two students qualify their commentary on the complexity of learning in English, stating that it is more difficult at the beginning and comparing the experience to learning in Spanish, making an evaluation that learning in English is more difficult. In Example 2, the students make an assessment of their future ability to speak English well, to which one AfL student responds negatively based on the difficulty of the language in general. Finally, in Example 3, the students comment on their level of anxiety in the CLIL classroom, with one AfL student citing more nervousness in classes taught in English than those taught in Spanish. Based on these responses, it seems that to these AfL students, English is quite complex, which makes them doubt whether they will be able to speak well in the future.

In the case of non-AfL students, fewer instances of negative complexity are found in relation to learning in English. However, some refer to test anxiety.

Extract 8.12: Non-AfL complexity in relation to test anxiety

Q: *Do you get more nervous when you have to take an exam in English than when you have to take an exam in Spanish?*
Non-AfL STU 1: *Well, a little more nervous. Because the questions in English, it's harder to know what you have to put*
Non-AfL STU 3*: Mmm, if it's in English, yes. Because how I look at it wrong, the verbs and I look at it wrong to write, well, it's hard for me. So I'm like, "what should I write?" So it's difficult for me.*

In extract 8.12 two non-AfL students refer to the difficulty of taking tests in English compared to using Spanish. The first student claims that their level of nervousness is heightened when responding to questions in a foreign language. The second student's response is similar, giving further detail regarding the mental processes that take place when this nervousness is present. However, despite these strong feelings of stress in an exam situation, when these non-AfL students are asked if they believe that learning in English is difficult their answers are neutral, as extract 8.13 shows.

Extract 8.13: Non-AfL Neutral complexity

Q: *Do you think that learning in English is easy or difficult?*
Non-AfL STU 1: *Normal. Because some words (are) difficult*
Non-AfL STU 2*: It's not really easy or difficult either because the teacher explains it to you.*

The first non-AfL student answers that learning in English is of moderate difficulty, mainly due the vocabulary. The second student also gives a neutral response, explaining that the teacher gives enough support to students. These examples differ from AfL student responses, which claim that the complexity of the language element increases the challenges of learning in a bilingual environment.

8.4.3. Social Valuation

SOCIAL VALUATION is an evaluation of the relevance or usefulness of a particular thing or topic. This was the most commonly used form of evaluation in the category of APPRECIATION (see table 8.13), with some students giving opinions on the usefulness of their citizenship courses and the English language.

Table 8.13: Social Valuation

SOCIAL-VALUATION-TYPE	AfL		Non-AfL		
positive-social-valuation	9	14.75%	13	19.12%	0.43
negative-social-valuation	0	0.00%	0	0.00%	0.00
neutral-social-valuation	1	1.64%	3	4.41%	0.82

Most examples in this category were positive, with slightly more instances in the Non-AfL interviews (13/19.12%) than the AfL interviews (9/14.75%). Few examples of neutral SOCIAL VALUATION were found, with only one instance in AfL (1/1.64%) and three in non-AfL interviews (3/4.41%). Positive social valuation found in the AfL interviews refers to learning the English language and the benefits this will bring in the future.

Extract 8.14: AfL: Social valuation of the English language

Example 1
Q: *Do you like to learn in English?*
AfL STU 3: *Yes, because I like to learn in new languages. I don't know, because for when I'm older and they speak to me and I have an English friend and not be speaking in Spanish. If I go to England or London.*

Example 2
Q: *Do you prefer to learn in Spanish or in English?*
AfL STU 2*: In English. Because there are other places where they speak more English and to....to go to these places so that they will understand me.*

Example 3
Q: *Do you think that English will be useful to you in the future?*
AfL STU 1: *Yes, because if I go to a place ...well, and I ask someone where is, I don't know what...well they tell me, well they tell me, well, if there's someone in English, well, they will tell me in English and I will understand.*
AfL STU 2: *Uh, yes. Because in that way, eh, I can speak English and....I can do useful things.*
AfL STU 3: *Yes. Like to make new friends or to go to England or to some English restaurant.*

AfL students gave a variety of explanations regarding the usefulness of English, most of which were based on envisioning their future selves interacting with speakers of English for travel purposes. In Extract 8.14, Example 1, the student answers that learning English facilitates future travel and integration with other cultures. Example 2 (Extract 8.14) refers to using English for travel purposes, though this student is more concerned with communication and being understood when abroad. In Example 3 (Extract 8.14), students are asked about the usefulness of English in the future. All students respond positively to this question, stating that knowing English will be useful to them in a variety of different situations such as travelling, making new friends and going to public places (such as restaurants) in which English is spoken. Through these statements, AfL students acknowledge that while learning in English is a complex task, it will prove beneficial for them in the future.

The non-AfL group expressed a similar stance regarding the benefits of learning English and how it will positively impact their future.

Extract 8.15: Non-AfL positive social valuation of learning English

Example 1
Q: *Do you like to learn in English?*
Non-STU 1*: Yes, because sometime when you are going to travel to some place, if you speak English they understand you better.*

Example 2
Q: *Do you prefer to learn in Spanish or in English?*
Non-AfL STU 3*: In English. Because this way, when I'm older, if I travel to a place and they're not, for example, French...if you speak French, well, anyway, well, (indecipherable) English and in that way everyone will understand me if I ask for something.*

Example 3
Q: *Do you think that English will be useful to you in the future?*
Non-AfL STU 1*: Well, yes.*
Non-AfL STU 2: *Yes, for when I go to other countries and they will understand me*
Non-AfL STU 3: *Yes, because I want to work as a scientist and scientists do it in Spain and also in another country.*

Like their AfL counterparts, these non-AfL students express positive beliefs on the benefits of learning English and how it will serve them in the future, citing traveling and mutual understanding when speaking to people from different places. One answer cites a positive SOCIAL VALUATION related to instrumental motivation for learning English, explaining that it will help achieve a future career goal of being a scientist. Through these responses, the non-AfL students also demonstrate an understanding of the benefits of learning English and how it will help prepare them for the future.

8.4.4. Summary of appreciation

Analysis of students' use of APPRECIATION shows that more instances of REACTION and COMPOSITION were found in AfL student interviews, while SOCIAL VALUATION was found more in non-AfL interviews. The majority of this APPRECIATION focused on the students' perceptions of the difficulty and long-term value of their language learning process. Overall, AfL students held positive attitudes toward learning in English, citing the fact that it will provide opportunities for travel and integration in other cultures, yet also expressed the difficulty of English learning. The non-AfL students also recognized the value of learning English, though commented on the difficulty of being tested in a second language. Nevertheless, all students acknowledged the social value of the English language and the positive effects that it would have in the future.

8.5. Chapter summary

This chapter concluded the presentation of the findings of this dissertation by presenting results obtained through an analysis of interviews with six lower-achieving CLIL students. A comparison of the AfL and non-AfL student responses was discussed, highlighting students' use of APPRAISAL as a means of reflecting on their learning process, their attitudes toward learning English and their capacities as language learners. The chapter continued with an analysis of students' use of JUDGEMENT, giving examples from each of the six categories. APPRECIATION was then focused upon, giving an overview of the categories and elaborating on qualtitative

results by providing extracts on the students' views of the difficulties and future benefits of English language learning.

The findings show no differences in the frequency of APPRAISAL in the student interviews were found. When investigating instances in each category, the distribution of types was similar, possibly as a result of the structured format of the interviews. In JUDGEMENT, most instances were seen in the categories of CAPACITY and TENACITY, which students used to self-assess their English speaking abilities and drive to try hard during the lesson. Surprisingly, many students from both groups expressed self-criticism regarding their abilities to communicate in English. It seemed that AfL students articulated specific areas (pronunciation, comprehension) that presented the most difficulty while non-AfL student answers were more general. In the case of APPRECIATION, instances from the three categories were found with equal distribution. Students expressed a general appreciation for learning in English, as reflected in instances of QUALITY, with AfL students citing a potential for the increased learning of vocabulary. However, AfL students also cited the COMPLEXITY of learning in a foreign language, expressing worry regarding speaking English well in the future due to its complex nature. Instances of COMPLEXITY found in the non-AfL interviews centered on test anxiety, with students commenting on a greater degree of difficulty taking a test in English than in their native language. Finally, both groups of students used SOCIAL VALUATION to comment on the value of learning English to travel and make friends abroad. The next chapter provides a discussion of the findings from Chapters 6, 7 and 8.

Chapter 9 Discussion

9.1 Introduction

The purpose of this study was to determine the relationship between the use of assessment for learning practices and teacher L2 motivational strategies found in CLIL lessons through analysis of a classroom corpus. Additionally, the study examined the effects of AfL implementation on learners' self-reported motivation, and feelings during the lesson in a questionnaire and interviews with lower achieving students.

In this chapter, a discussion of the relevant findings is provided for each of the three research perspectives that form part of this dissertation. The organization of the chapter follows the presentation of the results and aims to address the following specific areas:

a) A comparison of the **frequency and distribution** of L2 motivational strategies found in AfL and non-AfL lessons through analysis of the classroom corpus.
b) A comparison of the **duration** of L2 motivational strategies found in AfL and non-AfL units by **subject** (citizenship, science, drama and art).
c) A discussion of AfL techniques used by CLIL teachers[18] and their relationship to the **L2 motivational strategies** found in the corpus[19].
d) **Students'[20] self-reported motivation** expressed based on the content of the lessons as well as their attitudes toward learning in English
e) **Self-reported feelings** of AfL and non-AfL students in specific classroom situations
f) Self-evaluation and reflections by six **lower achieving CLIL students**

This chapter synthesizes the results presented in Chapters 6, 7 and 8, providing a discussion on their significance and relevance to the literature. The chapter draws

[18] Four teachers total: two AfL and two non-AfL.
[19] The corpus was comprised of 14 recorded lessons with a total duration of 9 hours 57 minutes. Total word count: 71,504 words.
[20] Students from AfL (n=19) and non-AfL (n=21) citizenship classes.

conclusions based on this analysis in relation to assessment practices in the context of CLIL primary lessons.

9.2. The relationship between the use of AfL techniques by teachers and L2 motivational strategies

This section addresses the results found in the analysis of the L2 motivational strategies found (see Chapter 6) comparing AfL and non-AfL lessons. The research questions addressed are the following:

Research Question 1: Do the frequency and distribution of second language motivational strategies differ depending on the use of AfL?

Research Question 2: How does the duration of these L2 motivational strategies vary depending on the subject (science, citizenship, art, drama)?

Research Question 3: Are there any L2 motivational strategies found in AfL lessons that are not identified in non-AfL lessons?

Research Question 4: Is there a relationship between teachers' use of AfL techniques and L2 motivational strategies observed during CLIL lessons?

The presentation of the chapter mirrors the organization of the findings, beginning with a discussion of the results found when analyzing the classroom corpus. The differences in frequency and distribution of L2 motivational strategies between the AfL and non-AfL lessons are discussed first, followed by the real time duration of each strategy compared by academic subject. A reflection on the appearance of specific L2 motivational strategies and their relationship to AfL techniques implemented by the teacher is then presented.

9.2.1 Frequency and distribution of L2 motivational strategies

The findings indicate a higher frequency and more varied distribution of L2 motivational strategies in the classes with AfL trained teachers, as shown in table 9.1.

Table 9.1: MOLT frequency and distribution results

AfL	**Non-AfL**
Total frequency of L2 motivational strategies: **908**	Total frequency of L2 motivational strategies: **751**
Distribution: Significant differences -Stating the comunicative purpose (0.99%)++[21] -Referential questions (38.82%) -Effective praise (8.48%)+++ -Neutral feedback (3.85%) -Echo (11.45%) -Process feedback (0.66%)++ -Peer and self-correction (5.29%)+++ -Arousing curiosity and attention (2.20%)+++ -Personalization (6.17%)+++ -Promoting autonomy (4.30%) +++ -Pair work (1.10%)++ -Scaffolding (8.48%)++	**Distribution: Significant differences** -Stating the communicative purpose (0.13%) -Referential questions (48.08%)+++ -Effective praise (3.97%) -Neutral feedback (6.75%) +++ -Echo (16.95%)+++ -Process feedback (0.00%) -Peer and self-correction (0.40%) -Arousing curiosity and attention (0.40%) -Personalization (2.65%) -Promoting autonomy (1.32%) -Pair work (0.26%) -Scaffolding (5.70%)
Distribution: No significant difference Signposting (8.37%) Promoting cooperation (0.44%) Establishing relevance (2.53%) Group work (1.32%)	**Distribution: No significant difference** Signposting (8.87%) Promoting cooperation (0.66%) Establishing relevance (1.85%) Group work (0.66%)

The total frequency of strategies found in the classroom corpus was higher in AfL lessons (908) than non-AfL lessons (751). Additionally, as seen in table 9.1, the

[21] +++High significance
++ Medium significance

distribution found in AfL lessons was more balanced, meaning the AfL teachers used a wider variety of L2 motivational strategies during their lessons. Significant differences are seen in the distribution for 9/16 of the strategies in the AfL lessons compared to 3/16 in the non-AfL lessons. No significant differences in distribution were found in the categories of signposting, promoting cooperation, establishing relevance and group work.

The differences in distribution suggest that a wider variety of motivational exchanges and activities were put into practice by the AfL teachers. This was due in part to the implementation of AfL techniques. These techniques led to the inclusion of mediating artifacts, discussion-based learning, long peer and self-assessment sessions and giving students more autonomy by placing them in the role of active learners. Many AfL techniques corresponded to L2 motivational strategies, which is discussed further in section 9.2.4.

On the other hand, in non-AfL lessons, the majority of strategies were concentrated in the categories of referential questions and echo, which accounted for 65% of the total number of L2 motivational strategies found. The lack of variety in L2 motivational strategies was perhaps due to the ways in which the non-AfL classes were conducted, following a more traditional approach of lecture-style teaching or independent student work, with the exception of the non-AfL citizenship lessons.

9.2.2 Strategies found in AfL lessons that were infrequent or not present in non-AfL lessons

Two strategies in particular, elicitation of peer and self-correction and process feedback, were absent or found very infrequently in the non-AfL corpus. Peer and self-assessment is an important AfL technique that has also been found to support student motivation (Cauley and McMillan, 2010) and played a central role in the AfL units. The inclusion of peer and self-correction was probably the result of training in AfL, which encourages teachers to allow students to help each other take an active role in the assessment process (Clark, 2012). The inclusion of this technique led to many exchanges in which students evaluated their own work or the work of their peers with support from the teacher.

Process feedback was the only L2 motivational strategy found in the AfL units and not found in the non-AfL units. Process feedback is emphasized in AfL training, as it helps students point out gaps in the learning process and how to make improvements (Nicol and Macfarlane-Dick, 2006). Though this type of feedback was used infrequently, it played a role in AfL lessons, helping students reflect on learning protocol and giving suggestions for improvement. Therefore, the presence of peer and self-correction and process feedback in the AfL units seem to be another result of AfL training.

9.2.3 Differences in duration of L2 motivational strategies found by subject

Citizenship

The longest periods of time devoted to L2 motivational strategies were found in AfL and non-AfL citizenship units, as reflected in the percentage of the strategy duration compared to the duration of the lesson[22]. The percentage of time dedicated to L2 motivational strategies in AfL citizenship lessons was the highest of all the units. However, only six different strategies were found in the AfL and non-AfL citizenship units[23]. The strategy lasting for the longest duration in both units was group work. It seems that the topic and objective of the citizenship lessons, centering on promoting discussion and collaboration amongst students, influenced the use of motivational strategies by the teacher. Also, the topic of the units (AfL: emotions, non-AfL: democracy) facilitated the introduction of personalization. Nevertheless, while the duration of class time dedicated to L2 motivational strategies in citizenship units was the longest, the variety of strategies was not as diverse.

Science

In the case of science lessons, a greater percentage of class time was dedicated to L2 motivational strategies in the AfL unit[24]. The AfL science teacher also used a much wider variety of motivational strategies, many of which were not seen in the non-

[22] AfL citizenship: 71%/ non-AfL citizenship: 54%
[23] Six strategies used in citizenship lessons: signposting, referential questions, scaffolding, echo, peer and self-correction and group work.
[24] AfL science: 48%/ non-AfL science: 43%

AfL unit.[25] The reasoning for this is due in part to the numer of AfL techniques implemented, which put the students in a different role, and the structure of the lesson. The AfL teacher used an active learner approach in which students were instructed to seek answers to a number of research questions using the course material. The teacher also devoted a substantial portion of the unit to explaining the learning objectives, using mediating artifacts such as WALT and WILF and "I can statements" to make learning objectives clear so that students could research independently. Here, the strategy used for the longest duration during the AfL science unit was signposting (23 minutes), which coincided with the AfL technique of stating the learning aims.

Through incorporation of such techniques, which put the students in the position of active learners, the AfL teacher used a variety of L2 motivational strategies. These included: giving students research aims at the onset of the lesson (signposting), allowing students to work together (group work) and, with the use of classroom resources, to seek out the answers to questions with the teacher providing support (scaffolding) when necessary. The answers students obtained through their investigation were then presented to the class and students gave feedback to their peers on the validity of their responses (peer correction). The discussion-based format that students engaged in when responding to the research questions, both in pairs and with the whole class, has been identified as beneficial to primary school learners, especially in the case of science learning (Harlen, 2004). The teacher also aroused curiosity and attention in students by placing students into a more active role. These techniques align with Stark and Gray's proposal (2001), which is that a good classroom environment gives students the freedom to present their ideas, which stimulates confidence and promotes learning.

On the other hand, the non-AfL teacher used a more lecture-style approach, explaining a PowerPoint presentation on pre-history and ancient history and asking

[25] Strategies seen only in AfL science unit: establishing relevance, promoting curiosity, stating the communicative purpose, effective praise, peer and self-correction, group work

students to complete a worksheet based on the presentation in pairs. Pair work was the strategy used for longest duration (52 minutes), which accounted for 75% of the total duration of strategies used in non-AfL lessons. Few motivational strategies were found in the non-AfL science unit, which followed a more traditional teaching style with the teacher in the position of the authority and the students in a more passive role.

Drama/ Art

In the case of AfL drama and non-AfL art lessons, L2 motivational strategies were found for a shorter duration in general, with the longest duration lasting only five minutes. Nevertheless, the amount of time dedicated to L2 motivational strategies in the AfL drama unit[26] was nearly half of the total class time.

Like the AfL science lesson, the AfL drama class followed a more discussion-based approach, placing students into a more active role, and included many AfL techniques. The beginning of the unit focused on reviewing several drama objectives from the course (signposting, stating the communicative purpose) and creating improvised plays (group and pair work), which were later critiqued through peer assessment (eliciting peer and self-correction). The structure of the AfL drama lesson, interactive nature and incorporation of AfL techniques such as stating the learning aims and peer and self-correction resulted in a greater variety of strategies.

On the other hand, the unit in which the least amount of time was dedicated to using these strategies was the non-AfL art unit. In these lessons, students worked independently to complete a task. Though the teacher offered assistance (scaffolding) and guided students in their work through the use of questions (referential questions), the duration of strategies in general was quite low.

[26] AfL drama: 48%/ non-AfL art: 14%

Summary of frequency and distribution of L2 motivational strategies

In sum, the use of AfL techniques was connected to the appearance of certain L2 motivational strategies, leading to the students taking on a more active role in learning. The collaborative nature of the citizenship lessons included more group and pair work in the AfL and non-AfL units, resulting in a longer duration of strategies, though with less variety. On the other hand, through the incorporation of AfL techniques, the AfL science and drama teachers used a wider variety of strategies, accounting for nearly half of the total class time. The incorporation of AfL techniques such as stating the learning aims and peer and self-assessment also coincided with the L2 motivational strategies of signposting and eliciting peer and self-correction.

9.2.4 The relationship of AfL techniques and L2 motivational strategies

In reviewing the primary classroom corpus**,** several differences are visible in the teacher discourse and activities of the two groups, many of which could be attributed to the implementation of techniques found in AfL training. A discussion of L2 motivational strategies found and their connection to specific AfL techniques is given in the next sections.

Signposting and stating the communicative purpose

As mentioned in Chapter 3, the phase of initiating motivation is integral to providing a "state of cognitive or emotional arousal" (Williams and Burden, 1997). This initiation involves clarifying learning objectives and introducing a sense of continuity and purpose, included in the MOLT scheme as signposting and stating the communicative purpose. In the case of signposting[27] (stating the lesson objectives explicitly or connecting to a previous lesson), differences in frequency were not found between the AfL and non-AfL groups. However, the exchanges found in AfL lessons revealed a more in-depth account of the objectives, some based on long, discussion-based interactions between the teacher and students. These extended

[27] Signposting: AfL (76/ 8.37%) Non-AfL (67/8.87%)

sessions were often facilitated by mediating AfL artifacts, such as the WALT or WILF posters displayed by the AfL teacher, showing lesson objectives for the day. Sharing learning objectives has been noted as contributing to intrinsic motivation, as it directs students toward learning goals and enables teachers and students to monitor the learning process.

In the AfL science class, instances of signposting were found when the teacher gave extensive descriptions of learning criteria paired with a corresponding grade through the use of the AfL technique "I can" statements. While the emphasis of grading in AfL is diminished, the combination of formative and summative assessment can make students aware of the criteria necessary to achieve a certain goal, or grade. Reflection on the criteria and the process of mapping out a path to achieve such a goal has been found to increase self-efficacy (Schunk and Swartz, 1993).

Feedback

After learning objectives are clear, and specific, achievable goals have been set, the use of feedback and peer and self-assessment occurs to fill learning gaps (Harrison Drozdowskij and Westhead, 2001). The use of feedback is an important factor in motivating, as evidenced by the inclusion of four distinct feedback types on the original MOLT scheme. Feedback takes a formative approach when students are advised on how to complete a task more effectively, resulting in student improvement (Clark, 2011). In the MOLT, this formative approach can be seen as process feedback, focusing on what can be learned from mistakes and how improvements can be made. As mentioned previously, process feedback was the only motivational strategy not found in non-AfL lessons and was implemented very little in AfL classes[28]. Considering the importance of formative feedback in AfL training, it was expected that process feedback would appear more frequently in the corpus. Nevertheless, when implemented, it was used to explain the learning

[28] Process feedback: AfL (6/0.66%), non-AfL: (0/0.0%)

protocol and make suggestions for improvement, which is one of the main goals of AfL (Black and et al., 2003).

Further differences in the use of feedback between the two groups included a tendency in the data toward the use of effective praise by AfL teachers. Effective praise[29] in the AfL class focused on improvements that students were making in order to attain their goals (e.g. "you are going up and up and up!"). This practice is integral to the creation of a positive learning environment by valuing and rewarding students who learn from their mistakes (Kaplan and Maehr, 1999). Rather than criticizing, it has been argued that praising student performance while offering constructive suggestions for improvement is the most effective form of feedback (Nichol and Macfarlane-Dick, 2007). Effective praise was also seen in the non-AfL corpus when commenting on student effort in completing an activity. The emphasis on effort over innate ability when praising student work has been cited as an important distinction to make in the interest of increasing student motivation (Aberger, 2010), possibly enhancing self-regulation, which drives students to be autonomous and goal-driven in their learning.

Neutral feedback[30] and echo[31] also appeared in the corpus, with echo found to be the most frequently used feedback type in both AfL and non-AfL lessons. While these two feedback types served to validate students' statements, the formative purpose was not visible, as the feedback given did not indicate areas of improvement.

Referential questions

While referential questions were the most frequently found motivational strategy in general, they were used more in non-AfL lessons[32]. Considering the emphasis on effective questioning techniques in AfL training, it was expected that referential

[29] Effective praise: AfL (77/8.48%), non-AfL: (30/3.97%)
[30] Neutral feedback: AfL (35/3.86%), non-AfL (31/6.75%)
[31] Echo: AfL (104/11.45%), non-AfL (128/16.95%)
[32] Referential questions: AfL (298/ 32.82%), non-AfL (363/ 48.08%)

questions would be found more frequently in the AfL corpus, but this was not the case. When considering the question types, open referential questions were found with a similar frequency in both groups, while closed referential questions were found more frequently in the non-AfL lessons. Referential questions were often used at the beginning of the AfL unit to measure students' previous knowledge of the material. Use of diagnostic questions has been cited as integral to AfL practice (Black et al., 2003) as a way of measuring the students' grasp of the material and adjusting lessons plans accordingly. In AfL lessons, this inquiry was enhanced through follow-up questions to encourage elaboration on student answers that may be incomplete. Furthermore, a finding seen only in AfL lessons was the use of referential questions for evaluation purposes (e.g. "What do you have to do now?" "How can you go to the next level?"). This finding supports the view of Birenbaum et al. (2009) regarding formative assessment as a cycle of inquiry, requiring reflective dialogue and encouraging students to identify learning gaps.

Peer and self-correction

The most notable difference between the AfL and non-AfL corpus was the presence of elicitation of peer and self-correction[33], which appeared very frequently in all AfL subjects and very little in non-AfL subjects. AfL training emphasizes peer and self-assessment techniques, and data from the corpus show that AfL teachers put these techniques into practice in a number of ways. When guiding students in peer and self-assessment, AfL teachers used a variety of techniques such as two stars and a wish, thumbs up/ thumbs down and reflecting on "I can" statements. Noting the active participation of students during these exchanges, it became apparent that these sessions were habitual and previous training had been provided for the students. Without training, it is difficult to ensure the students are able to assess in a meaningful way (Bingham, Holbrook and Meyers, 2010).

[33] Elicitation of peer and self-correction: AfL (43/ 5.29%)/ non-AfL: (3/ 0.40%)

The three peer correction sessions found in the non-AfL corpus focused on written error correction, taking a summative assessment approach. Conversely, those in the AfL lessons were discussion-based, focusing on positive aspects of student work and indicating areas of improvement. Through these sessions, students were able to take an active role in assessing their classmates (Leitch et al., 2007). This was often guided by the criteria provided by the teacher at the onset of the lesson through AfL techniques such as WALT and WILF posters and "I can" statements. Thus, students were able to check progress through reflection and provide feedback to their peers, becoming more involved in the learning process. This has been cited as an advantage of peer and self-assessment and may lead to increased self-esteem (Lindsay and Clarke, 2001). Finally, through the use of individual whiteboards, the AfL science teacher was able to lead peer and self-correction sessions, asking students to evaluate the work of their classmates in relation to both language and content.

Promoting student engagement

Several forms of promoting student engagement were seen in the AfL and non-AfL lessons, though the use of these strategies was most common in the AfL science and drama lessons. In the science lessons, both teachers were able to arouse curiosity and attention[34] in students through the use of visualization, or asking students to imagine themselves in certain situations during the lessons. This was also seen as the AfL teacher placed the students in their roles as "science detectives," asking them to sing the Pink Panther song with the learning objectives at the onset of the unit.

The method of placing students into the role of active learners in the AfL science lesson led to discussion-based pair work in which students were responsible for discovering answers to questions. During these sessions, the teacher provided scaffolding[35], which was used to help students complete an activity by reflecting on

[34] Arousing curiosity and attention: AfL (20/2.20%), non-AfL (3/0.40%)
[35] Scaffolding: AfL (77/8.48%), non-AfL (43 5.70%)

the learning process. The structure of independent investigation with the teacher acting as a mediator has shown to be integral in helping students cross their zone of proximal development (ZPD) (Vgotsky, 1978). Student collaboration with the teacher as a mediator was also seen in the non-AfL debate structure. This structure promoted collaboration and was used to establish the relevance of democracy to society by putting students in a mock-democratic situation and tying the content to their experience of present Spanish society. Establishing relevance was also seen in AfL lessons [36] to connect the classroom content to students' previous knowledge or personal experience.

Students were also able to connect content to personal experience through personalization[37], which was mainly found in AfL and non-AfL citizenship units. However, it appears that personalization was more closely connected to the topic of the unit, as it occurred the most frequently in the AfL citizenship unit in which students were studying emotions. Therefore, it cannot be claimed that personalization was due specifically to the use of AfL techniques.

In general, AfL students were given more autonomy [38] in both the regulative and instructional activities. This autonomy was given when selecting partners and making decisions related to the content or flow of the lesson, such as choosing characters to interpret in the AfL drama lesson. The promotion of student autonomy has been noted as a trend related to the implementation of AfL techniques (Marshall and Drummond, 2006). However, no significant differences were seen in the strategy of promoting collaboration, which was not used frequency in the corpus [39]. The few instances found were used for a similar purpose: placing students in a collaborative situation in which they were able to work together, sometimes with the assignment of roles (writer, speaker) to increase participation and a sense of responsibility.

[36] Establishing relevance: AfL (23/2.5%), non-AfL (14/1.85%)
[37] Personalization AfL (56/6.17%), non-AfL (20/2.65%)
[38] Promoting autonomy: AfL (39/ 4.30%), non-AfL (10/1.32%)
[39] Promoting cooperation: AfL (4/0.4%), non-AfL (5/0.66%)

Promoting collaboration

The appearance of discussion-based group[40] and peer work[41] was more frequent in the AfL lessons. This suggests that students were given more opportunities to work independently, resulting in active participation and taking responsibility for their own learning. By promoting freer interaction structures, the teacher is "letting go" of the control of learning and instilling a sense of responsibility in students, leading to the creation of student autonomy. Group and pair work was used for a wider range of activities in the AfL units, such as preparing for the speaking sections of the PET exam, discussing research questions and sharing personal experiences related to emotions. In the non-AfL units, group and pair work were found, though the format was task-based rather than discussion-based. The exception occurred in the non-AfL citizenship unit with the debate, in which students worked together to discuss the pros and cons of wearing a uniform. Collaboration in the form of group work was cited as an enjoyable activity in student responses on the questionnaire[42], therefore its implementation might be beneficial in motivating students and providing a more positive learning experience.

9.3 AfL and non-AfL students' self-reported motivation

The discussion continues with the students' perspective, analyzing the results found in student self-reported motivation and feelings in classroom situations through the motivational questionnaire.

The research questions addressed as part of this discussion are the following:

Research Question 5: Can any relation be seen between the type of assessment used and student's self-reported motivation? Are students in AfL classes more or less motivated than their non-AfL peers?

Research Question 6: How do AfL and non-AfL students describe their feelings in the context of certain classroom situations?

[40] Group work: AfL (12/1.32%), non-AfL (5/0.66%)
[41] Pair work: AfL (10/1.10%), non-AfL (2/0.26%)
[42] "I like working in groups" (AfL: 100%/ non-AfL: 95% of students responded "yes")

Sections 9.3.1 addresses RQ5 regarding students' self-reported motivation and 9.3.2 considers RQ6.

9.3.1 Questionnaire Part 1: Student self-reported motivation

The findings obtained in the first part of the study indicate a relationship between AfL training and an emphasis on certain L2 motivational strategies, suggesting that professional development in AfL techniques may help CLIL teachers motivate and engage their students. However, when investigating the impact on students, no significant differences were found between the two groups for each of the five categories of motivation in Part 1 of the questionnaire. It is possible that these findings are influenced by the small sample of students (N=40) selected to participate. However, this number is not uncommon when conducting AfL research, as most studies performed have been small scale, taking place in the context of professional development (Flórez and Sammons, 2013). The small sample of students was chosen in order to collect a variety of data to complement the analysis of the classroom corpus. Students from citizenship classes were chosen as the AfL citizenship teacher had more background with AfL training due to obtaining teacher education in the UK. However, while the AfL teacher did implement certain aspects of AfL training, the use of L2 motivational strategies in the citizenship lessons was the least varied when compared to other AfL units. Also, motivational strategies in the non-AfL citizenship unit were found frequently, which could account for the lack of difference in the students' self-reported motivation: both teachers used a similar variety of L2 motivational strategies and both groups of students appeared to be equally motivated.

Upon analysis of the individual items from the student questionnaire, AfL students demonstrated a higher level of uncertainty in their responses. In the category of self-efficacy, AfL students expressed uncertainty in their own abilities compared to their non-AfL counterparts. Specifically, a high percentage of students were unsure of whether others understood them as they were speaking English[43] and if they

understood as much as their classmates[44]. This finding contradicts literature claiming AfL techniques empower students by implementing mastery goals and self-assessment, thus giving higher expectations for success (Cauley and McMillan, 2010; Schunk, 1996). Doubts were also expressed in the categories of intrinsic value on the questionnaire, in which AfL students were uncertain of the usefulness of English, and self-regulation in which more AfL students expressed uncertainty of whether or not they tried hard during the lesson.

These findings can be interpreted in two different ways. For example, it could be that these specific AfL students demonstrate a greater sense of uncertainty and lower belief in their ability than their non-AfL peers, though no evidence was found in the teacher's behavior that might lead them to have lower confidence. The second interpretation is, due to extensive training in self-assessment, AfL students are better able to assess their own abilities and/or respond in a more honest way than their non-AfL peers. If the first possibility is true, further empirical research is necessary to explore the effects of AfL on students. Regarding the second possibility, it has been shown that extensive training is required to effectively help students become proficient in the self-assessment process (Ross, Siegenthaler and Tronson, 2006). Therefore, due to the process of constant questioning and reflection, AfL students may be more critical or realistic about their own abilities, accounting for the uncertainty in their responses. In this case, it would be recommendable for AfL teachers to focus on the positive learning outcomes of peer and self-assessment to facilitate progress and avoid excessive criticism.

The last finding regarding student self-reported motivation occurred in the category of test anxiety. In this case, AfL students demonstrated a low level of worry in a

---

[43] People understand me when I speak English:
AfL: **Yes:** 47% **I don't know:** 52% **No:** 0%
Non-AfL: **Yes:** 73% **I don't know:** 21% **No**: 5%

[44] I understand as much as my classmates:
AfL: **Yes:** 33%, **I don't know:** 47% **No:** 19%
Non-AfL: **Yes**: 58%, **I don't know**: 31% **No:** 10%

testing situation than their non-AfL counterparts[45]. Testing was less frequent in the AfL classrooms and did not comprise the main form of evaluation leading to the final grade, therefore alleviating testing pressure. In general, a de-emphasis on testing is a characteristic of AfL practice (Black et al., 2003) compared to a higher amount of pressure placed on testing in many classroom settings, which perhaps led to this result.

9.3.2. Questionnaire Part 2: Students' feelings about the class

Part 2 of the motivational questionnaire was designed to measure students' feelings in different classroom situations to determine whether a positive emotional environment was being created and possible effects on students' feelings and internal motivation. The importance of the classroom climate cannot be underestimated when giving students a forum for open expression of ideas and doubts without fear of ridicule (Hodgson and Pyle, 2010).

Two of the items on Part 2 of the questionnaire addressed classroom participation, specifically the feelings invoked when raising a hand to participate and when the teacher calls on a student to participate. While some students selected negative emotions such as 'timid' or 'silly' to describe their feelings when volunteering to participate, the overall response was mostly positive[46]. However, when commenting on their feelings when the teacher calls on them in class, the majority of non-AfL students selected negative emotions[47], with almost half indicating nervousness. On the other hand, the AfL response to this item was positive overall. This finding implies that classroom discussions with the teacher calling on students to participate was more frequent in the AfL lessons, and suggests that the climate was more welcoming when offering information. Perhaps the nervousness felt on the part of non-AfL students was a result of the fact this type of discussion, with the teacher asking students to participate spontaneously, was less habitual.

---

[45] I worry before an examination
AfL: **Yes:** 76% **I don't know:** 6% **No:** 19%
Non-AfL: **Yes:** 89% **I don't know:** 6% **No:** 5%
[46] When I raise my hand in class, I feel…(AfL +63%/ -36%, non-AfL +66%/-33%)
[47] When the teacher calls on me, I feel…(AfL +58%/-42%, non-AfL: +48%/-52%)

With regard to being helped by the teacher in class, the response for both groups was mixed[48], with half of the students responding positively with emotions such as 'confident', 'enthusiastic' and 'content and the other half saying they felt 'silly' or timid'. Accepting assistance may be construed as positive or negative based on the individual's viewpoint or the context, with some students viewing it as a sign of weakness. All students in both groups responded negatively when asked how they feel when they don't understand something in class. Despite the influence of AfL, which seeks to determine where learning gaps lie in order to fill these gaps, the feeling of not understanding still generates negative emotions.

When helping a classmate, all AfL students had positive feelings, while some non-AfL students reported feeling 'nervous'[49]. Group and peer work sessions were more habitual in AfL lessons, which may have influenced this result. Similarly, when being helped by a classmate, the majority of AfL students also gave positive responses while several non-AfL students cited negative emotions such as 'embarrassed', 'timid', and 'nervous'[50]. Asking a classmate for help was expected to provoke mainly negative responses due to the nature of expressing weakness or uncertainty, yet the majority of AfL students selected adjectives such as 'content' and 'confident', with few expressing a negative feeling ('silly'). This result could be the due to the use of peer and self-assessment, which was a common feature in AfL lessons. Therefore students were accustomed to working with their classmates and viewed providing constructive criticism positively. On the other hand, some non-AfL students who were not familiar with giving this type of structured feedback to peers expressed feeling 'embarrassed', 'timid' and 'nervous'.

Finally, when students commented on their feelings while taking an examination, more than half of AfL students reported feeling nervous during the examination[51]. This result contradicted the response in the first part of the questionnaire, in which

[48] When the teacher helps me, I feel… (AfL +47%/-53%, non-AfL: +53%/-47%)
[49] When I help a classmate, I feel…(AfL +100%/-0%, non-AfL: +91%/-9%)
[50] When a classmate helps me, I feel…(AfL +90%/-10%, non-AfL +77%/-23%)
[51] When I take an examination, I feel…(AfL +37%/-63%, non-AfL: +48%/-52%)

students were asked if they worry *before* an examination[52]. Possible explanation for this finding may be linked to less frequent testing in the AfL classroom, which means that students become nervous when finding themselves in a test taking situation. However, the AfL teachers' de-emphasis on the importance of grades and testing may explain the lack of worry prior to the examination.

9.4 Lower achieving students' reflections on their own learning

The final part of the study discusses results from interviews with lower achieving AfL and non-AfL students, focusing on the use of APPRAISAL when commenting on their learning process and feelings toward learning in English. The research question addressed is the following:

Research Question 7: How do lower achieving AfL and non-AfL students reflect on their own learning and classroom environment?

The findings indicate that lower achieving students in the AfL and non-AfL groups selected categories from certain APPRAISAL systems when commenting on their own learning experience, though no significant differences were found between groups. This finding may be due to the effects of the instrument, as students were encouraged to share their views in a structured way and were asked the same interview questions, prompting the same use of APPRAISAL. However, a detailed analysis revealed differences in the ways students used APPRAISAL to self-evaluate and comment on their learning experience, which is discussed in the next sections.

9.4.1 Judgement

Regarding JUDGEMENT, which evaluates human behavior, the most frequently used category was CAPACITY, as both groups of students provided self-evaluations of their own capabilities as learners and speakers of English. In this case, most students from both groups evaluated their English abilities in a negative way. However, the

[52] I worry before an examination
AfL: **Yes:** 76% **I don't know:** 6% **No:** 19%
Non-AfL: **Yes:** 89% **I don't know:** 6% **No:** 5%

ways in which these judgements were made differed between the AfL and non-AfL groups. Negative CAPACITY in non-AfL students was articulated in the sense of worry and lack of confidence in their English speaking abilities. One student specifically cited embarrassment that fellow students would laugh at their English during the class, in response to an interview question. Nevertheless, when non-AfL students were questioned regarding achieving high results in their classes taught in English, most gave positive responses regarding their own CAPACITY. This response suggests that, though students had doubts regarding their English level, they were unable to connect these doubts to possible consequences reflected in a low grade in the class. It may also suggest a lack of understanding of learning criteria, exhibited in the way non-AfL students commented on the challenges of learning in English, which was generalized without indicating specific areas of difficulty. Additionally, peer and self-assessment was found infrequently in the non-AfL corpus. This suggests that the lower achieving students interviewed were not accustomed to evaluating their own abilities based on learning criteria and so make connections to tangible results.

On the other hand, while AfL students also doubted their own abilities, specific doubts were cited relating to areas of language learning. For example, when asked about mistakes made during class, one AfL student described pronunciation difficulties by giving specific examples that created a cause for worry and doubt. Another AfL student expressed difficulty in comprehension, stating that at times they did not understand the teacher's instructions. These specific CAPACITY judgements may be attributed to the AfL students' peer and self- assessment training, which has been shown to identify areas of improvement (Lindsay and Clarke, 2001). This reflects the ability of AfL students to consider their own strengths and weaknesses and think about these difficulties in an objective way, though also reveals a lack of confidence that could prove to be detrimental.

Other findings show that students used TENACITY when describing the amount of effort expended throughout the unit. All students from the AfL group responded that they worked hard to make their teacher happy, while those from the non-AfL group were more neutral or negative. A possible explanation for this finding is based on

the continuous assessment administered during the AfL unit. This may provoke a sense of responsibility or accountability of students to their teachers, which is not seen in non-AfL lessons. Therefore, students feel less pressure to expend continuous effort to please the teacher. Another possible interpretation could be affectivity, or whether students were fond of their teacher, which may have led some to claim they work hard while others did not expend the same amount of effort.

9.4.2 Appreciation

In the case of APPRECIATION, students made use of all sub-categories with an even distribution. One particularly interesting finding occurred in the sub-category of QUALITY ("Do I like it?"), in which most AfL and non-AfL students expressed positive reactions toward learning in English. Two AfL students indicated that learning in a CLIL context meant that they "learn more", perhaps due to additional vocabulary in the L2, which coincides with findings on the benefits of CLIL (Dalton-Puffer, 2008). In the category of IMPACT ("Did it grab me") AfL students mainly commented on the content of their citizenship classes rather than language issues. Conversely, non-AfL lower achievers commented on learning in Spanish rather than English, suggesting a preference for learning content in their mother tongue rather than a foreign language.

Similarly, AfL students noted a greater level of difficulty due to the challenges of learning in English through use of COMPLEXITY. Despite expressing a positive attitude for CLIL, AfL student responses indicated a higher sense of anxiety regarding their language level and ability to speak English well in the future, due to its difficulty and complex nature. This finding coincided with the results regarding CAPACITY, in which AfL students expressed specific areas of doubt regarding their own abilities. As previously mentioned, two possibilities may explain this phenomenon, the first being a lower self-esteem or confidence in lower achieving AfL students, which may impact perceived competence. This is corroborated by a sense of uncertainty regarding language ability expressed by AfL students in the motivational questionnaire. The second possibility is a heightened awareness regarding the challenges involved in learning in an L2 due to exposure to AfL techniques, such as

peer and self-correction and a greater emphasis on explaining learning objectives and setting goals.

Regarding COMPLEXITY in relation to learning in English, non-AfL students were neutral, commenting that it was not easy or difficult. However, negative COMPLEXITY was found regarding test anxiety, specifically taking an examination in English. This finding coincided with results from the questionnaire indicating a higher percentage of non-AfL students worry before taking an examination in English. This may be attributed to the summative evaluation method in the non-AfL classroom, which caused heightened anxiety in these lower achieving students due to being tested in a foreign language.

One common theme in both groups was the value of learning English, which falls under the APPRAISAL category of SOCIAL VALUATION. Both AfL and non-AfL students recognized the importance of English learning for travel purposes, making new friends in a foreign country and being understood outside of Spain. This perspective coincides with Gardner's concept of integrativeness, or motivation to learn the material based on an interest in socializing with members from the target language community (Masgoret and Gardner, 2003). One non-AfL student also cited an instrumental factor related to working abroad as a scientist. These findings, while not related to AfL, show that, despite the perceived difficulty regarding the learning process, most AfL and non-AfL students were able to recognize the benefits of learning English for the future. This finding reflects the importance of the 4C's framework, (Coyle, 1999), which encourages CLIL teachers to emphasize the cultural dimension of language learning. In this case, this cultural aspect led learners to envision their future L2 self (e.g. traveling and speaking to foreigners when traveling abroad), as described in Dörnyei's L2 motivational Self System (Dörnyei 2005; Dörnyei, 2009).

9.5 Chapter summary

This chapter synthesized the findings in Chapters 6, 7 and 8, providing a discussion of emerging themes and their connection to AfL literature. The presence of L2

motivational strategies in the corpus was considered, comparing AfL and non-AfL lessons. A higher frequency and a more varied distribution of L2 motivational strategies were found in the classes with AfL trained teachers. It seems that the presence of AfL techniques caused a difference in the classroom, shifting away from traditional lecture-style teaching or individual work. This difference included more classroom discussion and group work with the teacher acting as a mediator, leading students to take a more active role in learning.

This active role was especially evident in AfL science lessons, in which a range of AfL techniques were implemented and a wide variety of L2 motivational strategies were found. One of the key factors was the active learner approach, which embodied the essence of AfL, encouraging students to take ownership of their own learning (Black and Wiliam, 2009), which leads to a greater sense of student motivation (Ryan and Deci, 2000). The most notable difference found between the AfL and non-AfL units was the presence of elicitation of peer and self-correction, a strategy seen highly frequently in all AfL subjects and very little in the non-AfL units. These exchanges added value to the lesson by providing students with a discussion-based forum to comment on their work and that of their peers, making them an active part of the assessment process.

Perhaps as a result of this technique, AfL students expressed a higher degree of uncertainty regarding their English level and understanding compared to that of their peers on the questionnaire. This might be related to the training AfL students had been exposed to when learning how to assess themselves and their peers in an effective way, heightening their metacognitive ability to be more self-critical when considering their own learning. This was especially visible in the interviews in which lower achieving AfL students commented on their CAPACITY, referencing specific aspects of language learning that proved to be challenging, such as pronunciation and comprehension and giving detailed descriptions of each. This was contrasted with their non-AfL peers, who gave much more general commentary.

In AfL and non-AfL citizenship lessons, both teachers dedicated a significant amount of time to these strategies, though the variety was not as diverse as in other units. Perhaps as a result of this, the motivational level of AfL and non-AfL citizenship students on the questionnaire did not show significant differences. However, when describing their feelings in the class, AfL students seemed to react more positively, even in potentially discouraging situations. This finding suggests a difference in the AfL classroom which led students to be more comfortable collaborating and helping each other fill learning gaps, possibly a result of training in peer and self-assessment practice.

A wide variety of AfL techniques in the drama unit led to more active participation of students through discussion, groups and pair work and participating in peer and self-correction. Conversely, the non-AfL art lesson contained the lowest amount of of time dedicated to L2 motivational strategies, perhaps due to the students working individually, which did not generate opportunity for discussion based work or interaction with peers.

Finally, when reflecting on the perspectives of lower achieving students in both AfL and non-AfL units, it appeared that the AfL students recognized the complexities of English learning, expressing doubts regarding their ability to speak English well in the future. This reaction coincides with the uncertainty regarding their CAPACITY, which, again, is possibly connected with training in peer and self-assessment, causing students to view their abilities and learning experience in a more critical way. Nevertheless, all students expressed positive reactions toward CLIL learning, recognizing the social value of learning a language, expressing views of their future selves speaking English when traveling and communicating with people abroad. While this finding does not suggest implications for AfL, it shows the added value of the culture aspect emphasized by CLIL (Coyle, 1999), which supports Gardner's socio-educational model. Chapter 10 draws conclusions based on these findings, providing suggestions for further research.

Chapter 10: Conclusion

10.1. Introduction

Chapter 9 discussed the findings of the three perspectives presented in this dissertation, which included analysis of AfL and non-AfL didactic units to identify L2 motivational strategies; AfL and non-AfL students' self-reported motivation and feelings during the lesson; and finally an examination of lower achieving students. This chapter presents conclusions, beginning with a summary of the study and a synthesis of the conclusions obtained based on the three perspectives analyzed. Suggestions for CLIL teachers are given, followed by limitations of the study. The chapter concludes with suggestions for further areas of research and final remarks.

10.2 Summary and review of the study

This study explored aspects of assessment in CLIL, considering the effects of AfL implementation on the teacher's use of motivational discourse in the classroom and the motivation of student learners. The motivation stemmed from the fact that assessment in CLIL is considered an emerging and under-researched area (Llinares, Morton and Whittaker, 2012; Coyle, 2010; Maggi, 2012; Barbero, 2012). At present, the issue of CLIL assessment is being addressed through the Assessment and Evaluation in CLIL (AECLIL) project funded by the European Commission, which documents assessment strategies currently being used in CLIL learning environments and proposes implementation of techniques that may be beneficial. However, empirical research is needed to determine the effects of these techniques on English language learners (Heritage et al., 2013; Alvarez et al., 2014) and lower achieving students, which was part of the purpose of this dissertation.

The literature on assessment was reviewed in Chapter 2, which set out to explore challenges of traditional assessment and describe how AfL can be used to complement summative assessment by addressing and filling learning gaps. Chapters 3 and 4 offered a review of the literature, focusing on motivation and Appraisal theory, which included the frameworks for analysis to be applied in the

data analysis in this study. Chapter 5 presented the methodology, with the primary objective of determining the relationship between AfL practice and L2 motivational strategies used during the lesson. This was achieved through analysis of a corpus of classroom discourse, which sought to identify L2 motivational strategies present in AfL and non-AfL lessons (Chapter 6). A second aim was to examine the self-reported motivation of groups of AfL and non-AfL students and also consider the feelings of these students in certain classroom situations (Chapter 7). These results were compared to the findings from the first part of the study to determine if any relationship could be seen between L2 motivation strategies in AfL and non-AfL lessons and the motivation and feelings of the students. Finally, the third perspective (Chapter 8) sought to determine the ways in which lower achieving students evaluated their own learning and classroom experience.

Within the limited scope of this study, it can be stated that the implementation of AfL coincided with an increase in L2 motivational strategies used by the teacher, which led to classes in which activities conducted were supported by a more motivational discourse. The distribution of motivational strategies was more varied in AfL units, meaning the teachers used a greater variety of strategies, incorporating them more frequently and for a longer duration. This was especially true in the case of AfL science and drama lessons, in which a wider range of strategies supporting student motivation were found. Analysis of the AfL units showed a greater focus on clarifying learning aims and criteria, and these were stated at the beginning and referred to consistently throughout the lesson. In AfL classes, use of feedback differed in that it centered on effective praise and included process feedback, while, in non-AfL classes, neutral feedback was more frequent. Process feedback was the only L2 motivational strategy found only in AfL lessons and, though used infrequently, helped guide students to meet learning goals. Peer and self-correction was used frequently in AfL lessons, and appeared very little in non-AfL units. Through peer and self-correction sessions, AfL teachers engaged students in dialogic exchanges (Kirton et al., 2007, Alexander, 2004) to evaluate language and content areas based on previously established criteria. Group and pair work were also found

more frequently in the AfL classroom, encouraging students to engage in discussion. Through these sessions, students were able to take on a more active role in the classroom. Finally, the active learner approach used in AfL science lessons was found to increase motivational strategies related to engagement, such as promoting autonomy and arousing curiosity and attention.

The second perspective of the study sought to measure the self-reported motivation of AfL and non-AfL students. Despite a higher frequency of L2 motivational strategies in AfL lessons, no significant differences were found in AfL and non-AfL students' self-reported motivation on the questionnaires. This finding was possibly due to the fact that the sample of AfL and non-AfL citizenship students who completed the questionnaire were both exposed to a similar variety and amount of strategies. Therefore, the motivation of both groups of students did not differ. However, it appeared that AfL students expressed more uncertainty regarding their own language abilities than their non-AfL peers. This finding was also seen in lower achieving AfL student interviews, in which these students used the APPRAISAL category CAPACITY to express a more critical view of their own capacity for learning in a foreign language and future abilities to speak English.

A possible explanation for these findings considered student exposure to peer and self-assessment, which might have made AfL students more critical of their abilities. Nevertheless, AfL students expressed positive feelings in classroom situations, even in potentially anxiety-provoking situations such as raising their hands, being called on by the teacher or being helped by a classmate. Other results included reports of lower test anxiety for AfL students, possibly due to a de-emphasis on the importance of testing in AfL classes. Finally, a positive SOCIAL VALUATION of the value of learning English was expressed by all lower achieving AfL and non-AfL learners.

The findings of this study contribute to an understanding of how the use of AfL techniques would make it possible to integrate motivational strategies in a more systematic way in the CLIL educational context. This is one of the first studies to

examine the effects of AfL in CLIL empirically using a corpus, and it is my hope that steps will be taken to continue this line of research.

In order to develop an understanding of AfL techniques and how they may lead to enhancing motivation, it is necessary to see how these techniques are applied in the classroom setting. This was the original intention of Black and Wiliam in the beginning stages of formative assessment development and research. At the onset, the two researchers drew the analogy of the classroom as a "black box", with a series of inputs (students, teachers, rules and requirements) and outputs (test results, more knowledgeable students and satisfied teachers) (Black and Wiliam, 1998b). The problem that Black and Wiliam identified was a lack of understanding of what was happening inside the "black box" when processing input to output. The basis for their research was to develop a clearer picture of classroom processes related to formative assessment through a literature review of relevant studies. It is my hope that this extensive analysis of the corpus, which included six video recorded units[53] with a total of 14 lessons, has helped gain more insight into the "black box" and to develop an understanding of the techniques, interactions and processes involved in AfL teaching and learning.

The study also highlights the necessity of contemplating the effects of AfL on the learner through the students' perspective, which has been absent in AfL research (Flórez and Sammons, 2013). This is especially true in the case of lower achieving students, a group that is perhaps the most affected by CLIL learning due to content and language-related demands. The need for research targeting this group of CLIL learners has been pointed out (Dobson, Pérez and Johnstone, 2010), which this dissertation sought to address by investigating the ability of these students to assess their own learning. Changes in educational policy frequently occur without consulting students directly to measure the positive or negative effects. This may sometimes lead to results that are not totally positive, for example, policy changes calling for high stakes testing (Stiggins, 2007, Black et al. 2002, Shohamy, 2001). In

[53] AfL & non-AfL citizenship; AfL & non-AfL science; AfL drama, non-AfL art.

the end, the students are the beneficiaries of these assessment techniques and the importance of seeking their perspective on issues related to the learning process is crucial.

10.3 Pedagogical implications for CLIL teachers

One of the intentions of the analysis conducted in this dissertation was to make teachers aware of specific AfL techniques that may be applied to produce a more motivational classroom discourse. When analyzing the corpus, it became apparent that a number of the AfL techniques shared the same purpose as L2 motivational strategies. For example, the importance of **stating the learning aims** has been emphasized in assessment for learning practice (Black and Wiliam, 1998a), leading to the development of several techniques that facilitate this purpose. The incorporation of WALT and WILF posters was seen at the beginning of almost every AfL session to make learners aware of learning aims, which corresponded to the L2 motivational strategy *signposting.* Through stating the learning aims and referring to them consistently, offering feedback to students regarding their progress, the teacher is able to generate goals and objectives for the students to strive toward. The use of "I can" statements served a similar purpose: identifying learning aims and providing specific reference points for achievement.

The integration of **peer and self-assessment** may be used to develop students' metacognitive awareness and make them more active learners and corresponds to the category *elicitation of peer and self-correction* on the MOLT scheme. Peer assessment emphasizes the role of students in the assessment process (Sadler, 1989). This approach contributes to making students critical of their learning process based on clearly established criteria. These interludes provide a moment of reflection, allowing students to determine whether learning objectives set at the beginning of the lesson have been met. The use of these AfL techniques was shown to be a valuable way of involving students in the lesson and providing feedback to be used when meeting learning goals.

This study also highlights the necessity of **placing students into the role of active learners with the teacher acting as a mediator for support**. Giving students autonomy means developing a sense of responsibility for their own learning. This helps students become independent thinkers, taking into account the learning aims and meeting them through scaffolding and support from their peers and the teacher. Following these active learning sessions, discussion may be completed to address relevant findings, with learners sharing information and assessing the work of their peers. This technique can be beneficial in establishing both content-and language-related learning aims and working with students to ensure the achievement of these goals. All these techniques shared the purpose of giving students responsibility for their own learning, showing them where they needed to go and guiding them on their journey.

10.4 Limitations of the study

This study is based on a relatively small sample size of teachers (N=4) and students (N=132). While a much larger corpus was collected and transcribed for the study, the six didactic units (14 total lessons, 71,504 total words) analyzed here were selected since they allowed comparison by topic. The detailed, manual, real-time analysis of the six units, achieving an in-depth picture of AfL techniques in the classroom context, presented here could not have been carried out on the larger corpus. Though the practice of using small sample groups is not uncommon in AfL research (Flórez and Sammons, 2013) the implementation of larger scale studies is necessary to avoid making generalizations or overemphasizing the effects of AfL from data based on a limited number of teachers and students (Black and Wiliam, 2003). In addition, the research context included only bilingual schools in Madrid, Spain where AfL practices had recently been implemented. Measurement of the effects of such techniques requires further investigation as AfL becomes more widespread.

10.5 Further research

While the role of assessment in CLIL is currently being considered through the AECLIL project (Barbero, 2012), there is a need to explore the possibilities of how assessment for learning and formative assessment can be integrated to help achieve content-and language-related goals. This has been identified as one of the greatest challenges in CLIL, with formative assessment being offered as a possible solution to connect these two areas (Coyle, 2010). However, implementation of AfL has been limited in CLIL contexts thus far, and more research is needed to determine its utility for teachers and students. This may include specifically investigating how AfL techniques can be used in a formative way to address learning gaps detected in the areas of content and language and how this might impact the students' learning experience. This insight may be gained through interviews with CLIL teachers trained in AfL to determine how the incorporation of AfL techniques addresses gaps in language and content. A longitudinal approach would also be beneficial, tracking the AfL teachers' assessment practices and their influence on student learning and motivation. However, in this study, my purpose was to conduct an in-depth analysis of the classroom recordings to discover the relationship of AfL and motivational strategies and understand the students' reactions.

10.6 Concluding remarks

CLIL has proven to be successful as a bilingual educational initiative throughout Europe, though the model is relatively young and still in a process of development (Dalton-Puffer, 2008). Assessment is one of the newest lines of research being studied as a mediating device for content and language demands placed on CLIL learners. Due to these language and content related challenges, it is also possible that these students may require a higher degree of motivation than learners in a traditional school setting. While motivation is a complex and abstract construct, teachers and researchers nevertheless share the common goal of searching for ways to sustain motivation in learners. It is my hope that this study has contributed to an

understanding of the ways AfL may be used in CLIL contexts, and may lead to more empirical research on assessment in CLIL to promote motivation.

References

Aberger, S. (2010). Formative Assessment: a review of the relevant research. *Expediontary Learning*.

Alexander, R. (2004). Towards dialogic teaching: rethinking classroom talk. Cambridge: Dialogos UK Ltd.

Alvarez, L., Ananda, S., Walqui, A., Sato, E. and Rabinowitz, S. (2014). Focusing formative assessment on the needs of English language learners. *WestEd.*

American Educational Research Association (2000). *AERA Position Statement concerning high stakes testing in PreK-12 Education.* Washington D.C.

Ames, C. & Ames, R. (1984). Systems of student and teacher motivation: Toward a qualitative definition. *Journal of Educational Psychology*, *76*, 535-556.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of educational psychology*, *84*(3), 261-271.

Angelo, T.A. & Cross, P.K. (1993). *Classroom Assessment Techniques* (2nd ed.). San Francisco: Jossey-Bass.

Askew, M. & Wiliam, D. (1995). Recent Research in Mathematics Education 5-16 London: The Stationery Office

Assessment Reform Group (2002). Assessment for Learning: 10 Principles. Retrieved from: http://www.qca.org.uk/libraryAssets /media/4031_afl_principles.pdf.

Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, N.J.: Van Nostrand.

Barbero, T. (2012). Assessment tools and practices in CLIL. In Quartapelle, F. (Ed.), *Assessment and Evaluation in CLIL (pp.* 38- 56). Retrieved from: www.aeclil.net.

Barkovska, I. (2012). Nutrition. In Quartapelle, F. (Ed.), *Assessment and Evaluation in CLIL (pp.* 117-128). Retrieved from: www.aeclil.net.

Beaton, A. E., Mullis, I. V., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1997). *Mathematics achievement in the middle school years: IEA's Third*

International Mathematics and Science Study. Boston College: Center for the Study of Testing, Evaluation and Educational Policy.

Bennett , N. & Kell, J. (1989). *A good start? Four year olds in Infant schools.* Oxford: Blackwell.

Bergan, J.R., Sladeczck, I.E., Scaz, R.D. & Smith, A.N. (1991). Effects of a measurement and planning system on kindergartners' cognitive development and educational programming. *American Educational Research Journal*, *28*, 683-714.

Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, *70*(3), 489-520.

Biggs, J. (1998). Assessment and classroom learning: A role for summative assessment? *Assessment in Education, 5* (1), 103-110.

Bingham, G., Holbrook, T., & Meyers, L. E. (2010). Using self-assessments in elementary classrooms. *Phi Delta Kappan*, *91*(5), 59-61.

Birenbaum, M., Kimron, H., Shilton, H., & Shahaf-Barzilay, R. (2009). Cycles of inquiry: Formative assessment in service of learning in classrooms and in school-based professional communities. *Studies in Educational Evaluation*, *35*(4), 130-149.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in education*, *5*(1), 7-74.

Black, P. & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 139-148.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2002) Working inside the black box: assessment for learning in the classroom. London: King's College London, Department of Education and Professional Studies.

Black, P., Harrison, C., Lee, C., Marshall B. & Wiliam, D. (2003). *Assessment for Learning: Putting it into Practice*. New York, NY: Open University Press and McGraw-Hill.

Black, P. & Wiliam, D. (2003). In praise of educational research: Formative assessment. *British Educational Research Journal*, *29* (5) 623-637.

Black, P., Harrison, C., Hodgen, J., Marshall, M. & Serret, N. (2008). Validity in teachers' summative assessments. *Assessment in Education, 17*(2), 215-232.

Boston, C. (2002). The concept of formative assessment. *Practical Assessment, Research, & Evaluation*. *8*(9). Retrieved from http://PAREonline.net/getvn.asp?v=8&n=9.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on Formation and Summative Evaluation of Student Learning*. New York: McGraw-Hill.

British Council. Bilingual Education in Spain. Retrieved from: http://www.britishcouncil.es/en/partnerships/success-stories/bilingual

Broadfoot, P., Osborn, M., Panel, C. & Pollard, A. (1996). Assessment in French primary schools. *The Curriculum Journal*, *7*(2), 227-246.

Broadfoot, P. M., Daugherty, R., Gardner, J., Harlen, W., James, M., & Stobart, G. (2002). Assessment for learning: 10 principles. Cambridge, UK: University of Cambridge School of Education.

Brown, S., & McIntyre, D. (1993). *Making sense of teaching*. Buckingham: Open University Press.

Brown, J. D. (1988). *Understanding research in second language learning: A teacher's guide to statistics and research design*. Cambridge: Cambridge University Press.

Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System*, *39*, 523-532.

Butler, R. (1988). Enhancing and undermining intrinsic motivation; the effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, *58*, 1-14.

Butler, D.L. & Wira, P.H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research*, *65*, 245-281.

Cauley, K. M., & McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *83*(1), 1-6.

Carless, D. (2005). Prospects for the Implementation of Assessment for Learning. *Assessment in Education, 12* (1), 39-54.

Christie, F. (1995). Pedagogic discourse in the primary school. *Linguistics and education, 7*(3), 221-242.

Christie, F. (2002). *Classroom Discourse Analysis: A Functional Perspective*. New York: Continuum.

Chu, C.P.Y. (2014). Supporting new arrival students' engagement with picture books: analysis of teacher talk using the appraisal theory. Functional Linguistics, *1*(1), 1-17.

Cid, E., Granena, G., & Tragant, E. (2009). Constructing and validating the foreign language attitudes and goals survey (FLAGS). System, *37*(3), 496-513.

Clark, I. (2011). Formative assessment and motivation: Theories and themes. *Prime Research on Education (PRE),* 1(2), 27-36.

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, *24*(2), 205-249.

Clegg, J. (2012). Preface. In Quartapelle, Franca (ed.) *Assessment and Evaluation in CLIL (*pp. 5-7). Retrieved from: www.aeclil.net.

Clément, R., & Kruidenier, B. (1983). Orientations in Second Language Acquisition: I. The Effects of Ethnicity, Milieu, and Target Language on Their Emergence. *Language Learning*, (33), 273-291.

Coffin, C. (1997). Constructing and giving value to the past: An investigation into secondary school history. *Genre and institutions: Social processes in the workplace and school*. London: Cassell. 196-230.

Coffin, C. (2002). The voices of history: Theorizing the interpersonal semantics of historical discourses. *The Open University, 22*(4), 503-528.

Commission of the European Communities (1995). White Paper: *Teaching and Learning- Towards the Learning Society*, Objective IV. Council of Europe, Brussels, DGV

Consejería de Educación, Comunidad de Madrid. Bilingual schools in the Comunidad de Madrid. Retrieved from: www.madrid.org.

Cooper, D. (2006). *Talk About Assessment: Strategies and Tools to Improve Learning*. Toronto: Thomas Nelson.

Covington, M.V. (1998). *The Will to Learn: A Self-Worth Perspective on Motivation and School Reform*. Cambridge: Cambridge University Press.

Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual review of psychology*, *51*(1), 171-200.

Coyle, D. (1999). Theory and planning for effective classrooms: supporting students in content and language integrated learning contexts. In Masih, J (Ed.): *Learning Through a Foreign Language (*pp. 1-26). London: CILT.

Coyle, D. (2008). CLIL – a pedagogical approach. In N. Van Deusen-Scholl, & N. Hornberger (Eds.), *Encyclopedia of Language and Education, 2nd edition* (pp. 97-111). Berlin: Springer Science+Business Media LLC.

Coyle, D. (2010). Assessment Issues in CLIL. In Coyle, D., Hood, P. and Marsh, D. (Eds.), *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press.

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL Content and Language Integrated Learning.* Cambridge: Cambridge University Press.

Crooks, T. (1988). The Impact of Classroom Evaluation Practices on Students. *Review of Educational Research*. *58* (4). 438-481.

Crookes, G., & Schmidt R. W. (1991). Motivation: Reopening the Research Agenda. *Language Learning*, *41*(4), 469-512.

Cross, R., & Gearson, M. (2013). Research and evaluation of the content and language integrated learning (CLIL) approach to teaching and learning languages in Victorian schools. Melbourne Graduate School of Education. Retrieved from http://www.education.vic.gov.au/school/principals/curriculum/Pages/clilvictoria. aspx.

Cumming, J. & Maxwell, G.S. (1999). Contextualising authentic assessment. Assessment in Education: Principles, Policy and Practice, *6*(2), 177-194.

Csizér, K. & Dörnyei, D. (2005). Language Learners' Motivational Profiles and Their Motivated Learning Behavior. *Language Learning*. *55*(4). 613-659.

Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated Learning (CLIL) Classrooms*. Amsterdam: John Benjamins.

Dalton-Puffer, C. (2008). Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe. In Werner Delanoy and Laurenz Volkmann, (Eds.), *Future Perspectives for English Language Teaching* (pp. 1-19). Heidelberg: Carl Winter.

Deci, E. L., Nezlek, J., & Sheinman, L. (1981). Characteristics of the rewarder and intrinsic motivation of the rewardee. *Journal of personality and social psychology*, *40*(1), 1-10.

Deci, E. L., & Ryan, R. M. (1992). The initiation and regulation of intrinsically motivated learning and achievement. *Achievement and motivation: A social-developmental perspective*, *3,* 9-36.

Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of educational research*, *71*(1), 1-27.

Dekkers, A., Howard, P., Adams, N., & Martin, F. (2014). Strategies to remove barriers and increase motivation to use the Tablet PC in formative assessment. *Using Technology Tools to Innovate Assessment, Reporting, and Teaching Practices in Engineering Education*, 164-177.

Denman, J., Tanner, R. & de Graaff, R. (2013). CLIL in junior vocational secondary education: challenges and opportunities for teaching and learning. *International Journal of Bilingual Education and Bilingualism, 16*(3), 285-300.

Derewianka, B. (2007). Using appraisal theory to track interpersonaldevelopment in adolescent academic writing. In McCabe, A., O'Donnell, M, and Whittaker, R. (Eds.), *Advances in language and education* (pp. 142-165). London: Continuum.

Diener, C.I. & Dweck, C. S. (1978). An analysis of learned helplessness: Continuous changes in performance, strategy, and achievement cognitions following failure. *Journal of Personality and Social Psychology*, *36*(5), 451-462.

Dietel, R.J., Herman, J.L., & Knuth, R.A. (1991). What does research say about assessment. *NCREL, Oak Brook.* Retrieved from: http://www. ncrel. org/sdrs/areas/stw_esys/4assess. htm.

Dobson, A., Pérez Murillo, M.D. & Johnstone, R. (2010). Bilingual Education Project: Evaluation Report. Ministry of Education (Spain), British Council.

Doiz, A., Lasagabaster, D., & Sierra, J.M. (2014). CLIL and motivation: the effect of individual and contextual variables. *The Language Learning Journal. 42* (2). 209-224.

Dörnyei, Z. (1990). Conceptualising motivation in foreign-language learning. *Language Learning, 40*, 45-78.

Dörnyei, Z. (1994). Motivation and Motivating in the Foreign Language Classroom. *The Modern Language Journal*, *78* (3), 273-284.

Dörnyei, Z., & Csizér, K. (1998). Ten Commandments for Motivating Language Learners: Results of an Empirical Study. *Language Teaching Research, 2*(3), 203-229.

Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching*, *31*, 117-135.

Dörnyei, Z. (2001). *Motivational strategies in the foreign language classroom*. Cambridge: Cambridge University Press.

Dörnyei, Z., & Csizér, K. (2002). Some dynamics of language attitudes and motivation: Results of a longitudinal nationwide survey. *Applied linguistics*, *23*(4), 421-462.

Dörnyei, Z. (2003). Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. In Z. Dörnyei (Ed.), *Attitudes, orientations and motivations in language learning* (pp. 3-32). Oxford: Blackwell.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z. (2009). *The L2 Motivational Self System*. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self (*pp. 9-42). Bristol: Multilingual Matters.

Dunn, K. E. & Mulvenon, S. W. (2009). A Critical Review of Research on Formative Assessments: The Limited Scientific Evidence of the Impact of

Formative Assessments in Education. Practical Assessment Research & Evaluation, *14*(7).

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological review*, *95*(2), 256.

The Eberly Center for Teacher Excellence. Why and How to conduct assessment. Retrieved from: http://www.cmu.edu/teaching/assessment/howto/.

Elliott, E. S., & Dweck, C. S. (1988). Goals: an approach to motivation and achievement. *Journal of personality and social psychology*, *54*(1), 5-12.

Eurydice (2006). Content and Language Integrated Learning (CLIL) at School in Europe. Brussels: Eurydice.

European Commission (2006). Europeans and their languages. *Special Eurobarometer.* 243. Brussels: European Commission.

Flórez, M. T., & Sammons, P. (2013). *Assessment for Learning: Effects and Impact*. Evaluative Report. CfBT Education Trust, University of Oxford (England), Department of Education.

Fairclough, N. (1992). *Discourse and social change*. Cambridge: Polity Press.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist, 34*(10), 906-911.

Fontana, D. (1996). Changes in control beliefs in Portuguese primary school pupils as a consequence of the employment of self-assessment strategies. *British Journal of Educational Psychology, 66*, 301-313.

Fontana, D., & Fernandes, M. (1994). Improvements in mathematics performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology*, *64*(3), 407-417.

Furtak, E. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. Applied Measurement in Education, *21*(4), 335-359.

Gardner, R. C. & Lambert, W.E. (1959). Motivational Variables in Second-Language Acquisition. *Canadian Journal of Psychology, 13*, 266-272.

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and Motivation in Second-Language Learning*. Rowley, MA: Newbury House.

Gardner, R. C. (1982). Language attitudes and language learning. In E. Boudhard Ryan & H. Giles*, Attitudes towards language variation* (pp. 132-147). London: Edward Arnold.

Gardner, R. C. (1985). *Social Psychology and Second Language Learning: The Roles of Attitudes and Motivation*. London: Sage Publications.

Gardner, R. C., & Tremblay, P. F. (1994). On Motivation, Research Agendas, and Theoretical Frameworks. *The Modern Language Journal*, *78*(3), 359-368.

Gardner, R. C. (2001). Language Learning Motivation: The Student, the Teacher, and the Researcher. *Texas Papers in Foreign Language Education*, *6*(1), 1-18.

Gardner, R.C. (2007). The importance of motivation in second language acquisition. *Porta Linguarum, 8*, 9-20.

Gardner, R. C. (2010). *Motivation and second language acquisition.* New York: Peter Lang.

Garrison, C., & Ehringhaus, M. (2007). *Formative and Summative Assessments in the Classroom*. Dover, NH: Measured Progress.

Gobierno de España, Ministerio de la Presidencia (1990). Ley Orgánica de Ordenación General del Sistema Educativo. Retrieved from: http://www.boe.es/buscar/doc.php?id=BOE-A- 1990-24172.

Great Britain Office for Standards in Education (1998). Secondary education 1993-1997: A review of secondary schools in England.

Guilloteaux, M. J. & Dörnyei, Z. (2008). Motivating language learners: A classroom oriented investigation of the effects of motivational Strategies on student motivation. *TESOL Quarterly*, *42*(1), 55-77.

Guion, L. A., Diehl, D. C., & McDonald, D. (2002). Triangulation: Establishing the validity of qualitative studies. University of Florida IFAS Extension. Retrieved from http://edis.ifas.ufl.edu/fy394

Halbach, A. (2009). The primary school teacher and the challenges of bilingual education. In Dafouz, E. and M. Guerrini (Eds.), *CLIL across educational levels: experiences from primary, secondary and tertiary contexts* (pp. 19-26). Madrid: Santillana.

Halliday, M.A.K. (1993). Towards a language-based theory of learning. *Linguistics and education, 5*(2), 93-116.

Halliday, M.A.K. (1994). *Introduction to Functional Grammar*, 2nd edition, London: Edward Arnold.

Halliday, M.A.K., and Matthiessen C.M.I.M. (2004). *Introduction to Functional Grammar,* 3rd edition, London: Edward Arnold.

Hargreaves, A. (1989). *Curriculum and Assessment Reform.* Toronto: OISE Press.

Harlen, W., & Deakin Crick, R. (2002). A systematic review of the impact of summative assessment and tests on students' motivation for learning. *Research evidence in education library*, Vol. *1.*

Harlen, W. (2004). Talking and writing: have we got the balance right? *Primary Science Review*, *83*, 17–19.

Harlen, W. (2007). *Assessment Of Learning.* London: Sage Publications.

Harmon, M. (1995). The changing role of assessment in evaluating science education reform. In O'Sullivan, R.G. (Ed). Emerging roles of evaluation in reforming science education (special issue) *New Directions for Program Evaluation, 65*, 31-52.

Harrison, C., Drozdowskij, J. & Westhead, K. (2001). Formative assessment in primary classrooms, Primary Science Review, *68*, 19–22.

Harter, S. (1981). The development of competence motivation in the mastery of cognitive and physical skills: Is there still a place for joy? In G.C. Roberts & D.M. Landers (Eds.), *Psychology of motor behavior and sport-1980*. Champaign, IL: Human Kinetics.

Heritage, M., Walqui, A., & Linquanti, R. (2013). Formative assessment as contingent teaching and learning: perspectives on assessment as and for language learning in the content areas. *American Educational Research Association meeting, San Francisco, CA*.

Heyman, G. D., & Dweck, C. S. (1992). Achievement goals and intrinsic motivation: Their relation and their role in adaptive motivation. *Motivation and emotion*, *16*(3), 231-247.

Higgins, E. T. (1987). Self-discrepancy: a theory relating self and affect. *Psychological review*, *94*(3), 319-340.

Hodgson, C., & Pyle, K. (2010). *A literature review of Assessment for Learning in science*. Slough: National Foundation for Educational Research.

Hönig, I. (2010). *Assessment in CLIL: theoretical and empirical research*. Saarbrücken, Germany: VDM Verlag Dr. Muller.

Hyland, K. (1996). Writing without conviction: hedging in science research articles. Applied Linguistics, *17* (4), 433–54.

Iedema, R., Feez, S., & White, P. R. (1994). Media literacy (Write it right literacy in industry research project–Stage 2). Sydney: Metropolitan East Disadvantaged Schools Program.

Johnson, R. B., Onwuegbuzie, A.J., & Turner, L.A. (2007). Toward a definition of mixed methods research. *Journal of mixed methods research 1*(2), 112-133.

Johnson, D. W. & Johnson, R. T. (1999). Learning Together and Learning Alone: Cooperative, Comparative and Individualistic Learning. University of California: Allyn and Bacon.

Kaplan, A., & Maehr, M. L. (1999). Achievement goals and student well-being. *Contemporary educational psychology*, *24*(4), 330-358.

Kellaghan, T., Madaus, G. F. & Raczek, A. (1996). *The Use of External Examinations to Improve Student Motivation.* Washington, D.C.: American Educational Research Association.

Kellard, K., Costello, M., Godfrey, D., Griffiths, E., & Rees, C. (2008). Evaluation of the Developing Thinking and Assessment for Learning development programme. *Welsh Assembly Government*.

Kiely, R. (2009). CLIL– The question of assessment. Retrieved from: developing-teachers. com.

Kirton, A., Hallam, S., Peffers, J., Robertson, P. & Stobart, G. (2007). Revolution, evolution or a trojan horse? Piloting assessment for learning in some Scottish primary schools*. British Educational Research Journal*, *33*(4), 605-627.

Lasagabaster, D. (2008). Foreign Language Competence in Content and Language Integrated Learning Courses. *The Open Applied Linguistics Journal*, Vol. I, 30-41.

Lasabagaster, D. (2011). English achievement and student motivation in CLIL and EFL settings. *Innovation in Language Learning and Teaching*, *5*(1), 3-18.

Lasagabaster, D. & López Beloqui, R. (2015). The Impact of Type of Approach (CLIL Versus EFL) and Methodology (Book-Based Versus Project Work) on Motivation. *Porta Linguarum, 23*, 41-57.

Leitch, R., Gardner, J., Mitchell, S., Lundy, L., Odena, O., Galanouli, D., & Clough, P. (2007). Consulting pupils in Assessment for Learning classrooms: the twists and turns of working with students as co-researchers. *Educational Action Research*, 15(3), 459-478.

Lemke, J. L. (1992). Interpersonal meaning in discourse: value orientations. In M. Davies & L. Ravelli (Eds.), *Advances in Systemic Linguistics: Recent Theory and Practice* (pp. 82–194). London: Pinter (Open Linguistics Series).

Lindsay, C. and Clarke, S. (2001). Enhancing primary science through self– and paired–assessment. *Primary Science Review*, *68*, 15–18.

Llinares, A., Morton, T., & Whittaker, R. (2012). *The roles of language in CLIL*. Cambridge University Press.

Lorenzo, F., Casal, S, Moore, P., Afonso, T. (2009). *Bilingüismo y educación. Situación de la Red de Centros bilingües en Andalucía*. Sevilla: Centro de estudios Andaluces. Consejería de la Presidencia, Junta de Andalucía.

MacIntyre, P. D., MacKinnon, S. P., & Clément, R. (2009). The baby, the bathwater, and the future of language learning motivation research. In Dörnyei, Z., & Ushioda, E. (Eds.), *Motivation, language identity and the L2 self* (pp. 43-65) Bristol: Multilingual Matters.

Maggi, F. (2012). Evaluation in CLIL. In Quartapelle, F. (Ed.), *Assessment and Evaluation in CLIL* (pp. 57-75). Retrieved from: www.aeclil.net.

Maljers, A., Marsh, D., Coyle, D., Hartiala, A., Marsland, B., & Perez-Vidal, C. (2001). *CLIL: Compendium*. Retrieved from: http://www.clilcompendium.com.

Markus, H., & Nurius, P. (1986). Possible selves. *American psychologist*, *41*(9), 954-969.

Marshall, B., & Drummond, M.J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research papers in education, 21*(2), 133-149.

Marsh, D. (2002). *CLIL/EMILE- The European Dimension: Actions, Trends and Foresight Potential* Public Services Contract DG EAC: European Commission.

Marsh, D. (2006). Special Educational Needs in Europe: The Teaching and Learning of Languages. European Commission: Public Services Contract Brussels: European Commission.

Martin, J. R. (1995). Interpersonal meaning, persuasion and public discourse: packing semiotic punch. *Australian Journal of Linguistics, 15*(1), 33–67.

Martin, J. R. & Rose, D. (2003). *Working with discourse: Meaning beyond the clause, Second Edition*. London: Continuum.

Martin, J.R. & White, P.R.R. (2005). *The Language of Evaluation: Appraisal in English.* Hampshire: Palgrave MacMillan.

Martinez, J.G.R., & Martinez, N.C. (1992). Re-examining repeated testing and teacher effects in a remedial mathematics course, British Journal of Educational Psychology, *62*, 356-363.

Masgoret, A. M., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language learning*, *53*(S1), 167-210.

McCallum, B., Hargreaves, E. & Gipps, C. (2000). Learning: the pupil's voice. *Cambridge Journal of Education, 30*(2), 275-289.

McCabe, A. & Whittaker, R. (Forthcoming). Genre and appraisal in CLIL history texts: developing the voice of the historian. In Morton, T. & Llinares, A. (Eds.), *Applied Linguistics Perspectives on CLIL*. Amsterdam: John Benjamins.

Meyer, O. (2010). Towards quality-CLIL: successful planning and teaching strategies/ Hacia un CLIL/AICLE de calidad: estrategias efectivas de planificación y enseñanza. PLUSO. *33*. 11-29.

Mills, K. A., Unsworth, L., Bellocchi, A., Park, J.Y., & Ritchie, S. M. (2014). Children's Multimodal Appraisal of Places: Walking with the Camera. *Australian Journal of Language and Literacy, 37*(3), 171-181.

Natrillo, G. (1987). The impact of evaluation processes on students. *Educational Psychologist. 22*(2), 155-175.

National Council of Teachers of English (2014). How standardized tests shape- and limit-student learning. A policy research brief produced by the National Council of Teachers of English. Retrieved from: http://www.ncte.org/library/NCTEFiles/Resources/Journals/CC/0242-nov2014/CC0242PolicyStandardized.pdf

Nicol, D. J. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice, Studies in Higher Education, *31*(2), 198–218.

O'Donnell, M. (2008). The UAM CorpusTool: Software for corpus annotation and exploration. In Bretones Callejas, Carmen M. et al. (Eds.), *Applied Linguistics Now: Understanding Language and Mind / La Lingüística Aplicada Hoy: Comprendiendo el Lenguaje y la Mente* (pp. 1433-1447). Almería: Universidad de Almería.

O'Donnell, M. (2013). Exploring Identity through Appraisal Theory: a Corpus Annotation Methodology. *Linguistics and the Human Sciences*, *9*(1), 95-116.

Pascual Peña, I. (2010). Teachers' questions in CLIL contexts. *Vienna View(z*), *19*(3), 65-71.

Pérez-Cañado, M. L. (2012). CLIL research in Europe: past, present, and future. International Journal of Bilingual Education and Bilingualism, *15*(3), 315-341.

Pink, D. (2009). Drive: The Surprising truth about what motivates us. NY: Penguin Group.

Pintrich, P. & DeGroot, E. (1989). *Motivated Strategies for Learning Questionnaire.* University of Michigan.

Pollard, A. (1990). Towards a new educational perspective on children's learning. *Education 3-13, 11*(3), 241-256.

Popham, J.W. (2006). *Assessment for Educational leaders.* University of Michigan: Allyn and Bacon.

Poehner, M. E. & Lantolf, J. P. (2005). Dynamic Assessment in the Language Classroom. Language Teaching Research, *9*(2), 233-265.

Raynor, J. O., Atkinson, J. W., & Brown, M. (1974). Subjective aspects of achievement motivation immediately before an examination. In Atkinson, J.W. & Raynor, J.O. (Eds.), *Motivation and achievement* (pp. 155-171). Washington D.C.: Winston & Sons.

Rea-Dickins, P. (2001). Mirror, mirror on the wall: identifying processes of classroom assessment. *Language Testing, 18*(4), 429-462.

Read, J., Hope, D. & Carroll, J. (2007). Annotating Expressions of Appraisal in English. In *Proceedings of The Linguistic Annotation Workshop, ACL 2007*. Prague, Czech Republic. 93-100.

Reay, D. & Wiliam, D. (1999). I'll be nothing: structure, agency and the construction of identity through assessment. *British Educational Research Journal*, *25* (3), 343-354.

Reilly, T. & Medrano, P. (2009). MEC/ British Council Bilingual Project: Twelve years of education and a smooth transition to secondary. In Dafouz, E. and Guerrini, M. C. (Eds.), *CLIL Across Educational Levels* (pp. 59-70.). Madrid: Santillana Education, S.L./ Richmond Publishing.

Ross, P. M., Siegenthaler, S. L. & Tronson, D. A. (2006). Assessment for Learning and Motivation. UniServe Science Assessment Symposium Proceedings. 120-125.

Rowe, M. B. (1974). Reflections on wait-time: Some methodological questions.*Journal of Research in Science Teaching*, *11*(3), 263-279.

Rowe, J. & Coonan, M.C. (2008). Some foreign language issues in Primary CLIL: the teacher's voice. *International conference. The Future of Education. Pixel.* Milan.

Ruiz de Zarobe, Y. & Lasagabaster, D. (2010). Ways forward in CLIL: provision issues and future planning. In David Lasagabaster and Yolanda Ruiz de Zarobe

(Eds.), *CLIL in Spain: Implementation, Results and Teacher Training (pp.* 278–95). Newcastle: Cambridge Scholars Publishing.

Rumberger, R. W. & Thomas, S. L. (2000). The distribution of dropout and turnover Rates among urban and suburban high schools. *Sociology of Education, 73*(1), 39-67.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well being. *American psychologist*, *55*(1), 68.

Sadler, D. R. (1987). Specifying and Promulgating Achievement Standards. *Oxford Review of Education*, *13*, 191–209.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119-144.

Schunk, D. H., & Swartz, C. W. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology*, *18*(3), 337-354.

Schunk, D.H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. American Educational Research Journal, *33*, 359-382.

Seikkula-Leino, J. (2007). CLIL learning: Achievement levels and affective factors. *Language and Education*, *21*(4), 328-341.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational researcher*, *29*(7), 4-14.

Shohamy, Elana. (2001). *The Power of Tests: A Critical Perspective on the Uses of Language Tests.* London: Longman.

Spada, N., & Frolich, M. (1995). *COLT Observation Scheme: Communicative Orientation of Language Teaching Coding Conventions*. National Centre for English Language Teaching and Research.

Spiller, D. (2009). Assessment Matters: Self-Assessment and Peer Assessment. *Teaching Development Unit.* New Zealand: The University of Waikato.

Stark, R. and Gray, D. (2001). What science do they think they know? *Primary Science Review*, *68*, 8–11.

Stiggins, R.J., Griswold, M.M. & Wikelund, K.R. (1989). Measuring thinking skills through classroom assessment. *Journal of Educational Measurement*, *26*, 233-246.

Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, *86*(1), 22-27.

Stiggins, R. (2005). ¨From Formative Assessment to Assessment FOR Learning: A Path to Success in Standards-Based Schools¨, *Phi Delta Kappan*, *87*, 324-328.

Stiggins, R. (2007). Five Assessment Myths and their Consequences. *Education Week*, *27*(8), 28-29.

Sutton, R. (1995). *Assessment For Learning.* Oxford, England. RS Publishers.

Sylvén, L. K. & Thompson, A. (2015). Language learning motivation and CLIL: Is there a connection? *Journal of Immersion and Content-Based Language Education, 3*(1), 28-50.

To, J., & Carless, D. (2015). Making productive use of exemplars: Peer discussion and teacher guidance for positive transfer of strategies. *Journal of Further and Higher Education*, 1-19.

Townsend, T. (1997). What Makes Schools Effective? A Comparison between School Communities in Australia and the USA. *School Effectiveness and School Improvement*, *8*(3), 311-326.

Tunstall, P. & Gipps, C. (1996). Teacher Feedback to Young Children in Formative Assessment: A Typology. *British Educational Journal*, *22*(4), 389-404.

UK Department for Education (2013). Primary assessment and accountability under the new National Curriculum. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/298568/Primary_assessment_and_accountability_under_the_new_curriculum_consultation_document.pdf.

US Department of Education. (2001). The Elementary and Secondary Education Act (The No Child Left Behind Act of 2001). Retrieved from: http://www2.ed.gov/policy/elsec/leg/esea02/index.html.

Wall, K. (2008). Understanding Metagocognition Through the Use of Pupil Views Templates: Pupil Views of Learning to Learn. *Thinking Skills and Creativity, 3*, 23-33.

Webb, M. & Jones, J. (2009). Exploring Tensions in Developing Assessment for Learning. Assessment in Education: Principles, Policy and Practice. *16*(2), 165-184.

Weiner, B. (1992). Human motivation: Metaphors, theories and research. Newbury Park, CA: Sage

Weiner, B. (2000). Attributional thoughts about consumer behavior. *Journal of Consumer research*, *27*(3), 382-387.

Wewer, T. (2014). Assessment of Young Learners' English Proficiency in Bilingual Content Instruction CLIL. (Doctoral dissertation). Retrieved from: http://www.doria.fi/bitstream/handle/10024/96838/AnnalesB385Wewer.pdf?sequence=2

Whiting, B., Van Burgh, J. W., & Render, G F. (1995). Mastery learning in the classroom. Paper presented at the Annual Meeting of the American Educational Research. Association, San Francisco, CA.

Wiggins, G. (1998). *Educative Assessment*. San Francisco. CA: Jossey Publishing.

Wiliams, M. & Burden, R. L. (1997). *Psychology for language teachers*. Cambridge: Cambridge University Press. 111-141.

Williams, M., Burden, R. L., & Al-Baharna, S. (2001). The role of the individual in motivation theory. *Motivation and second language acquisition*, *23*, 171-184.

Wilson, M. (2009). Assessment from the Ground Up. *Phi Delta Kappan*, *91*, 68-71.

Ushioda, E. (2001). Language learning at university: Exploring the role of motivational thinking. In Dörnyei, Z. and Schmidt, R., (Eds.), *Motivation and second language acquisition* (pp. 93-125). Honolulu, HI: University of Hawaii, Second Language Teaching and Curriculum Center.

Ushioda, E. & Dörnyei, Z., (2011). Motivation, language identity and the L2 self: A Theoretical Overview. In Dörnyei, Z., & Ushioda, E. (Eds.), *Motivation, language identity and the L2 self* (pp. 1-9). Bristol: Multilingual Matters.

Vygotsky, L. (1978). Interaction between Learning and Development. *Mind and Society.* Cambridge, MA: Harvard University Press. 79-91.

Yeh, S. S. (2009). Class size reduction or rapid formative assessment?: A comparison of cost-effectiveness. *Educational Research Review*, *4*(1), 7-15.

Yin, Y., Shavelson, R., Ayala, C., Ruiz-Primo, M.A., Brandon, P., & Furtak, E. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. Applied Measurement in Education, *21*(4), 335-359.

Young, E. (2005). Assessment for Learning: Embedding and extending. Assessment is for Learning.

Appendices

Appendix 1
Parental consent form

C.E.I.P "SAN SEBASTIAN" El Boalo / Cerceda / Mataelpino
http://www.educa.madrid.org/web/cp.sansebastian.elboalo

Estimados padres y madres de 5º,

Nuestro centro se encuentra entre los primeros en desarrollar el proyecto bilingüe en la Comunidad de Madrid. El nivel de competencia en lengua extranjera adquirido por nuestros/as alumnos/as es alto y este hecho hace que diferentes universidades se interesen por la manera en que los/as niños/as se comunican en el aula.
Este es el caso de una de nuestras Auxiliares de Conversación quién se encuentra cursando el doctorado en la Universidad Autónoma de Madrid y requiere nuestra colaboración para grabar alguna clase en la que participan vuestros/as hijos/as. Dichas grabaciones **no se publicarán** en ningún tipo de medio y únicamente se utilizarán dentro del ámbito de la investigación.
De acuerdo con la ley de protección de datos, os solicitamos la autorización para realizar dicha grabación dentro del aula.
Recibid un cordial saludo,
El equipo de maestros/as de 5º.

Autorizo a mi hijo/a _____ a que aparezca en grabaciones de video/audio con el objetivo de analizar la comunicación dentro del aula.

Firmado,

_____

Appendix 2
AfL techniques used by teachers

Stating Communicative Purpose	**Drama (AfL Teacher 1):** Teacher used WALT and WILF to state objectives at beginning and as she changed the activity continually reminded the students of what they were doing and why. **Citizenship (AfL Teacher 1):** Teacher asks students for a retroactive summary of what they have been discussing in class for the past few weeks and uses WALT and WILF posters to discuss the upcoming unit objectives. **Science (AfL Teacher 2):** Teacher uses WILF poster to state what she wants for them to achieve during the lesson. She also gives students "I can" statements with possible marks and what they must do in order to achieve this mark.
Effective Questioning Techniques	**Drama (AfL Teacher 1):** Teacher used effective questioning techniques throughout, asking them to remember what they had done during the year in drama and the important elements of acting and improvising. **Citizenship (AfL Teacher 1):** A large number or referential questions are used throughout the unit. **Science (AfL Teacher 2):** Teacher makes use of many referential questions throughout the lesson.
Feedback	**Drama (AfL Teacher 1):** Teacher gave students feedback on their performances and as they responded to questions. **Citizenship (AfL Teacher 1):** Teacher gives feedback to students as they discuss their emotions, neutral feedback and effective praised used throughout. **Science (AfL Teacher 2):** Lots of neutral feedback/ echo and some effective praise.
Peer Assessment	**Drama (AfL Teacher 1):** Teacher chose three students to give feedback as whole class improvised certain characters. They had to choose student who did the best job and why. Teacher also asked for

	feedback (two stars and a wish) when students improvised scenes. Their classmates had to tell them something that they did well and something they needed to improve. **Citizenship (AfL Teacher 1):** After each student has discussed the reasons they link a certain emotion with a certain color, teacher asks the class to comment on the reasoning and/ or applaud. **Science (AfL Teacher 2):** Teacher asks students to write down answers on whiteboards and then bring them up to the front. Students read their answer to each other, then teacher takes a few whiteboards and asks them to do "two stars and a wish" to provide critiques to their peers answers (grammar and content)
Self Assessment	**Drama (AfL Teacher 1):** Thumbs up/ Thumbs down. After students had to speak for a minute about a certain topic or improvised play, teacher would ask them if they found the activity easy or difficult (thumbs up/ thumbs down) **Citizenship (AfL Teacher 1):** Teacher asks students to assess the times that they felt a certain emotion and how they reacted at the time. **Science (AfL Teacher 2):** Teacher asks students to give themselves two marks (one color: green, yellow, red) and one number (1, 2 and 3) based on their participation and how much English they have spoken throughout the class. Teacher also does "thumbs up/ thumbs down" intermittently throughout lesson.

APPENDIX A

Extract from the MOLT Classroom Observation Scheme

Learners' motivated behavior			Eager volunteering (>1/3 of the class)					
			Engagement (>2/3 of the class)					
			Attention (>2/3 of the class)					
Teacher's motivational practice	Encouraging positive retrospective self-evaluation		Class applause					
			Effective praise					
			Elicitation of self/peer correction session					
			Process feedback session					
			Neutral feedback session					
	Generating, maintaining, and protecting situation-specific task motivation	*Activity design*	+ team competition					
			+ individual competition					
			+ tangible task product					
			+ intellectual challenge					
			+ creative/interesting/fantasy element					
			+ personalization					
			+ tangible reward					
		P.S.[a]	Group work					
			Pair work					
		Teacher discourse	Referential Questions					
			Promoting autonomy					
			Promoting cooperation					
			Scaffolding					
			Arousing curiosity or attention					
			Promoting instrumental values					
			Promoting integrative values					
			Establishing relevance					
			Stating communicative purpose/utility of activity					
			Signposting					
			Social chat (unrelated to the lesson)					
minutes				1	2	3	4	5

Note. [a]P.S. = Participation structure

Appendix 4

Sample coding and re-coding sheets adapted from MOLT observation scheme

Subject: Science 1

Total Time: 37:25

Ref Q's	Scaffolding	Neutral Feed	Signpost.	Personal.	Effective Praise	Peer and Self Asses.
0.5 1.0 1.9 0.7 2.8 0.7 2.1 1.0 1.1 1.6 0.9 0.8 0.7 0.5 1.8 0.8 6.8 0.6 4.4 3.3 2.2 5.1 4.9 0.7 5.4 1.7 1.0 1.4 1.4 0.7 1.3 2.3 0.6 1.3 1.5 1.9 3.0 1.2 0.7 1.6 4.1 0.9 2.6 2.3 0.7 1.2 1.5 1.5 2.2 1.8 0.6 0.5 1.2 0.5 1.6 0.8 1.7 2.2 1.1 1.7 2.3 0.3 0.7 0.6 1.5 0.8 1.5 0.4 0.8 5.4 0.9 0.5 0.6 0.5 0.8 0.9 0.7 **2:30.8**	3.2 5.1 4.1 5.3 2.7 5.4 17.0 1.8 2.6 16.0 3.4 4.5 4.3 2.4 2.0 2.4 0.7 23.9 **1:27.8**	0.6 0.6 1.2 1.0 0.7 1.5 0.6 1.1 0.4 0.6 0.3 1.8 0.7 0.8 1.1 **1:07.0**	9.4 1.8 3.2 3.0 4.6 0.8 1.4 **00:24.2**	5.0 10.5 5.0 **00:20:5**	0.4 0.4 5.0 1.3 6.4 1.7 1.7 1.8 5.6 0.3 0.9 3.0 4.5 1.4 4.6 0.6 5.4 4.3 0.4 **00:49.7**	10.4 11.3 4.9 19.3 23.7 12.1 52.6 32.6 20.7 1:17.9 **4:25.5**
Promote Aut.	Arousing Cur.	Establishing Rel.	Stating Comm. Purp.	Process Feedback	Group Work	Pair Work
7.4 4.5 7.0 2.4 **00:21.3**	3.9 2.5 21.8 4.4 **00:32.6**	17.9 2.4 **00:20.3**	31.0 11.0 16.8 2.8 53.7 13.8 **2:09.1**	**0:00.0**	**0:00.0**	40.0 1:51.9 **1:31.9**

Class Structure

Teacher greets class and asks students what they need to do to get organized: Homework and open agenda on desk, whiteboard, marker, book and the "boxes to tick". Teacher asks students to respond to question: "Is a candle a natural or artificial light source?" on their whiteboards. Students then come to front of class and share their answers with one another, and then teacher takes two whiteboards and asks students to offer two pieces of positive feedback and one piece of criticism. After, teacher reads WILF to students and "I can" statements to give students an idea of what they need to master as the unit progresses. Teacher then asks students to write the grade they will be trying to get on their whiteboards and sing "I know you want me, you know I want you" to their board. Students then take out their science book and materials. Teacher writes several questions on the board ("What is a definition of sound? " "How is sound produced?") and gives students opportunity to find the answers with a partner before engaging in whole class discussion. They then go over the homework for that evening (Write two sentences about questions on blackboard) and then teacher asks students to self-assess their participation/ English during the class (a color, green yellow or red and a number, 1, 2, or 3).

Appendix 4

Sample coding and re-coding sheets adapted from MOLT observation scheme

Subject: Science 1

Total Time: 37:25

Ref Q's	Scaffolding	Neutral Feed	Signpost.	Personal.	Effective Praise	Peer and Self Asses.
-3.1 -4.5 -1.0 -1.0 -0.7 -1.1 -1.1 +2.3 -5.5 +2.3	-2.2 -1.3	-1.3 -0.5 -0.4 -1.3 -0.6 -0.6 -1.2 -1.3	-5.6 +18.4 +30.2 +10.0 +2.3 +50.5 -1.0 -0.8 +11.3		-0.6 -0.3 -0.3 -0.3 -0.3 -0.3	-1.9 +14.3
Promote Aut.	Arousing Cur.	Establishing Rel.	Stating Comm. Purp.	Process Feedback	Group Work	Echo
+6.0 +6.4			-18.4 -30.2 -10.0 -2.3 -50.48 -11.3		Pair	+2.1 +1.2 +1.3 +0.6 +0.5 +0.4 +1.1 +0.4 +1.9 +0.4 +0.6 +0.3 +0.4 +1.3 +0.6 +0.6 +1.2 +1.2 +0.5 +1.3 +0.7 +0.6

Motivated Strategies for Learning Questionnaire*

Please rate the following items based on your behavior in this class. Your rating should be on a 7-
point scale where 1= not at all true of me to 7=very true of me .

1. I prefer class work that is challenging so I can learn new things.
2. Compared with other students in this class I expect to do well
3. I am so nervous during a test that I cannot remember facts I have learned
4. It is important for me to learn what is being taught in this class
5. I like what I am learning in this class
6. I'm certain I can understand the ideas taught in this course
7. I think I will be able to use what I learn in this class in other classes
8. I expect to do very well in this class
9. Compared with others in this class, I think I'm a good student
10. I often choose paper topics I will learn something from even if they require more work
11. I am sure I can do an excellent job on the problems and tasks assigned for this class
12. I have an uneasy, upset feeling when I take a test
13. I think I will receive a good grade in this class
14. Even when I do poorly on a test I try to learn from my mistakes
15. I think that what I am learning in this class is useful for me to know
16. My study skills are excellent compared with others in this class
17. I think that what we are learning in this class is interesting
18. Compared with other students in this class I think I know a great deal about the subject
19. I know that I will be able to learn the material for this class
20. I worry a great deal about tests
21. Understanding this subject is important to me
22. When I take a test I think about how poorly I am doing
23. When I study for a test, I try to put together the information from class and from the book
24. When I do homework, I try to remember what the teacher said in class so I can answer the questions correctly
25. I ask myself questions to make sure I know the material I have been studying
26. It is hard for me to decide what the main ideas are in what I read
27. When work is hard I either give up or study only the easy parts
28. When I study I put important ideas into my own words
29. I always try to understand what the teacher is saying even if it doesn't make sense.
30. When I study for a test I try to remember as many facts as I can
31. When studying, I copy my notes over to help me remember material
32. I work on practice exercises and answer end of chapter questions even when I don't have to
33. Even when study materials are dull and uninteresting, I keep working until I finish
34. When I study for a test I practice saying the important facts over and over to

myself
35. Before I begin studying I think about the things I will need to do to learn
36. I use what I have learned from old homework assignments and the textbook to do new assignments
37. I often find that I have been reading for class but don't know what it is all about.
38. I find that when the teacher is talking I think of other things and don't really listen to what is being said
39. When I am studying a topic, I try to make everything fit together
40. When I'm reading I stop once in a while and go over what I have read
41. When I read materials for this class, I say the words over and over to myself to help me remember
42. I outline the chapters in my book to help me study
43. I work hard to get a good grade even when I don't like a class
44. When reading I try to connect the things I am reading about with what I already know.

*Pintrich, R. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance, Journal of Educational Psychology, 82, 33-40.

Appendix 5

Original and adapted Motivated Strategies for Learning Questionnaire

Nombre: _____ Profesor/a: _____

Asignatura: _____ Fecha: _____

1ª Parte : Por favor, responde con una de las siguiente palabras:

En esta unidad……

1. Levanté la mano y participé	SÍ	A VECES	NO
2. Me he esforzado mucho	SÍ	A VECES	NO
3. He hecho todos los deberes y actividades	SÍ	A VECES	NO
4. Pedí ayuda cuando no entiendía	SÍ	A VECES	NO

2ª Parte: Por favor, elije una respuesta

5. Espero sacar una buena nota	NO	SÍ	NO SÉ
6. Me gustó el tema de la clase	NO	SÍ	NO SÉ
7. Creo que he hecho bien todos los deberes	NO	SÍ	NO SÉ
8. Entiendo tanto como mis compañeros	NO	SÍ	NO SÉ
9. Me entienden cuando hablo inglés	NO	SÍ	NO SÉ
10. Me preocupo mucho antes de un examen	NO	SÍ	NO SÉ
11. Saber inglés es muy útil	NO	SÍ	NO SÉ
12. Me gusta participar en clase	NO	SÍ	NO SÉ
13. Me gusta trabajar en grupos	NO	SÍ	NO SÉ

3ª Parte:
Por favor, escribe un adjetivo para describir cómo te sientas en estas situaciones.
Ejemplos: Contento, seguro, emocionado, inteligente
Triste, torpe, nervioso, vergonzoso, tímido, frustrado

10. Cuando levanto mi mano en clase, me siento _____

11. Cuando mi profesor/a me llama para contestar, me siento _____

12. Cuando mi compañero/a me ayuda, me siento _____

13. Cuando ayudo a mi compañero/a , me siento _____

14. Cuando mi profesor/a me ayuda, me siento _____

15. Cuando no entiendo algo, me siento _____

16. Cuando hago un examen, me siento _____

¡Gracias por responder a mi cuestionario!

Student Motivational State Questionnaire

Attitudes Toward the Course

· I wish we had more English lessons at school this semester.
· I like English lessons this semester.
· English is one of my favorite subjects at school this semester.
· When the English lesson ends, I often wish it could continue.
· I want to work hard in English lessons to make my teacher happy.
· I enjoy my English lessons this semester because what we do is neither too hard nor too easy.
· I would rather spend time on subjects other than English. (REVERSED)
· Learning English at school is a burden for me this semester. (REVERSED)
· In English lessons this semester, we are learning things that will be useful in the future.

Linguistic Self-Confidence

· I feel I am making progress in English this semester.
· I believe I will receive good grades in English this semester.
· I often experience a feeling of success in my English lessons this semester.
· I am sure that 1 day I will be able to speak English.
· In English lessons this semester, I usually understand what to do and how to do it.
· This semester, I think I am good at learning English.
· I am worried about my ability to do well in English this semester. (REVERSED)
· I often volunteer to do speaking presentations in English lessons.

L2-Classroom Anxiety

· I get very worried if I make mistakes during English lessons this semester.
· I am afraid that my classmates will laugh at me when I have to speak in English lessons.
· I feel more nervous in English class this semester than in my other classes.

Adapted Student Motivational State Questionnaire
Interview questions for lower achieving students

Attitudes toward the course:
¿A ti te gustan las clases de ciudadanía?
¿Prefieres ciudadanía que otras asignaturas?
¿Piensas que ciudadanía es fácil o difícil?
¿Cuando la clase de ciudadanía termine, quieres que continua?
¿Te esfuerzas en ciudadanía para que tu profe este contento/a?
¿Crees que los temas de ciudadanía serán útil en el futuro?
¿A ti te gusta aprender en inglés?
¿Prefieres aprender en español o inglés?
¿Crees que aprender en inglés es muy difícil?
¿Piensas que el inglés te va a servir en el futuro?

Linguistic Self-Confidence
¿Crees que estás mejorando este trimestre en tus clases?
¿Piensas que vas a sacar buenas notas en tus clases que están en Ingles este trimestre?
¿Te sientes que tienes éxito en tus clases que están en inglés?
¿Crees que un día, vas a hablar muy bien el inglés?
¿En tus clases que imparten en inglés, normalmente entiendes lo que tienes que hacer y como tienes que hacerlo?
¿Este trimestre, te sientas como estas aprendiendo mucho inglés?
¿Estas preocupado que este trimestre, no vas a tener éxito en tus clases que están impartidos en inglés?
¿Levantas la mano para hablar en clases de inglés, ciencias, y ciudadanía?

Anxiety in the L2 Classroom
¿Te preocupes si te equivoques cuando estas hablando en inglés?
¿Tienes miedo que tus compañeros se van a reír de ti cuando tienes que hablar en inglés en clase?
¿Te sientes mas nervioso/a en tus clases que están impartidos en inglés que en las otras clases?

Test Anxiety
¿A ti te gusta tomar exámenes? ¿Por que?
¿Te pones mas nervioso cuando tienes que tomar un examen en Ingles que un examen en español?
¿Estudiando por un examen te ayuda aprender más?
¿Estudiando por un examen te ayuda mejorar tu inglés? ¿Por qué?
¿Que actividad en particular te ayuda mejorar tu inglés?
¿Prefieres tomar exámenes o hacer actividades en clase? ¿Que actividades prefieres?

<table>
<tr><td>Judgement</td><td colspan="3">Evaluates human behavior ethically (morally and legally). Language which criticizes or praises, condemns or applauds the behavior - actions, deeds, sayings, motivations, etc. – of human individuals or groups.</td></tr>
<tr><td></td><td>Normality</td><td>Assessment of how special or unusual the behavior or the person’s state is e.g. neutral, familiar, fashionable, celebrated vs. artificial, poseur, eccentric, peculiar, odd.</td><td rowspan="2">Inscribed- The text contains a word that explicitly describes the behavior. E.g. Immoral, virtuous, just, sinful, lavacious</td></tr>
<tr><td></td><td>Capacity</td><td>Assessments of competence or ability e.g. skilled, knowledgeable, brilliant vs. stupid, dull, ignorant, clumsy</td></tr>
<tr><td></td><td>Tenacity</td><td>Assessments of psychological disposition with determine and resolve e.g. plucky, brave, resolute, reliable, loyal, hard-working vs. cowardly, reckless, hasty, impatient.</td><td rowspan="3">Invoked- The judgement is not explicit in a word or phrase, but rather implied by information given. E.g. Bush delivered his inaugural speech as the US president who</td></tr>
<tr><td></td><td>Propriety</td><td>Assessments of ethical or moral standing, ‘how far beyond reproach’ is the behavior or person’s state e.g. good, fair, just, generous, charitable, kind versus immoral, corrupt, arrogant, greedy,</td></tr>
<tr><td></td><td>Veracity</td><td>Assessments regarding the person’s truthfulness or honesty, dependent on social contextual</td></tr>
</table>

		values e.g. credible, candid, direct, sincere versus deceitful, a liar manipulative, devious	collected 537,000 fewer votes than his opponent
	Unclear	Assessments of behavior and persona where none of the sub-types of judgement may apply. These cases are different than those which need to be double coded under conditional of ambiguity	
Appreciation	Evaluates things, processes and states of affairs aesthetically or with the social value accorded to the object. Human participants may be 'appreciated' where the assessment does not directly focus on the correctness or incorrectness of behavior, but rather the appearance of the person		
	Reaction	Values which make reference to, or are derived from, values of affect but where the emotional reaction (depress, bore etc) has been detached from any human experiencer of the emotion and been attached to the evaluated entity as if it were some property which the entity objectively and intrinsically possesses e.g. A depressing sight met our eyes.	**Impact**: How does it strike me? What initial reaction does it make? the perceptual aspects or aesthetics of the item **Quality**: Do I like it? how do I react emotionally towards it? What are my affectual responses?
	Composition	How well the parts of the entity fit together. E.g. harmonious, well-formed, balanced, unified, intricate versus ill-formed, convoluted, confused, unbalanced,	**Balance**: Did it hang together? Was it harmonious, organized, well-proportioned, logical, or unbalanced, lop-sided, irregular, flawed, discordant, shapeless?

		discordant, contorted.	
			Complexity Was it hard or easy to follow? Was it simple, pure, elegant, clear, precise, lucid, coherent, or was it extravagant, byzantine, woolly, arcane, simplistic, etc?
	Social Valuation	Whether something is 'socially' valued for its usefulness, worthiness, efficaciousness, health-giving properties: its contribution to the community, or its value to the consumer: related to judgement.	

Appendix 8

Table of duration, mean and range for each L2 motivational strategy

	AfL Citizenship	Non-AfL Citizenship	AfL Science	Non-AfL Science	AfL Drama	Non-AfL Art
Signposting	0:18-1:00 Mean: 0:18 Sum: 5:00*	0:02-0:42 Mean: 0:13 Sum: 9:52*	0:01-12:54 Mean: 0:27 Sum: 22:40*	0:02-1:24 Mean: 18.1 Sum: 5:08	0:06-3:11 Mean: 0:27 Sum: 5:01*	0-0:22 Mean: 0:11 Sum: 1:08*
Referential questions	0.003-0:12 Mean: 0:02 Sum: 2:01	0.003s-0:10 Mean: 0:01 Sum: 3:47	0.003-0:06 Mean: 0:01 Sum: 5:23	0.3-0:12 Mean: 0:02 Sum: 5:56	0:004-0:05 Mean: 0:01 Sum: 0:41	0:003-0:04 Mean: 0:01 Sum: 1:33
Scaffolding	0.008-0:21 Mean: 0:06 Sum: 2:31	0:013-0:32 Mean: 0:09 Sum: 3:37	0:008-1:58 Mean: 0:11 Sum: 10:25*	0:03-0:21 Mean: 0:10 Sum: 1:35	0-0:10 Mean: 0:10 Sum: 0:10	0:03-1:53 Mean: 0:45 Sum: 6:00*
Establishing relevance	0:06-0:21 Mean: 0:12 Sum: 1:13	0-0:25 Mean: 0:11 Sum: 2:35	0:01-0:33 Mean: 0:12 Sum: 2:46*	0-0 Mean: 0.0 Sum: 0.0*	0-0 Mean: 0.0 Sum: 0.0	0-0 Mean: 0.0 Sum: 0.0
Promoting autonomy	0:02-0:14 Mean: 0:06 Sum: 0:41	0-0:07 Mean: 0:04 Sum: 0:34	0:006-0:21 Mean: 0:04 Sum: 1:42	0:009-0:009 Mean: 0:009 Sum: 0:009	0:02-0:09 Mean: 0:04 Sum: 0:36	0:01-0:06 Mean: 0:03 Sum: 0:09
Arousing curiosity or attention	0-0:04 Mean: 0:03 Sum: 0:06	0-0:38 Mean: 0:01 Sum: 0:41	0:006-2:42 Mean: 0:17 Sum: 4:38*	0:04-0:04 Mean: 0:04 Sum: 0:04	Range: 0:03-0:03 Mean: 0:03 Sum: 0:03	Range: 0-0 Mean: 0.0 Sum: 0.0
Stating the communicative purpose	0:01-0:18 Mean: 0:08 Sum: 1:02	0-0:14 Mean: 0:14 Sum: 0:14	0-0:50 Mean: 0:12 Sum: 2:13*	0-0 Mean: 0.0 Sum: 0.0	Range: 0:06-0:32 Mean: 0:21 Sum: 1:04*	Range: 0-0 Mean: 0.0 Sum: 0.0
Promoting cooperation	0-0 Mean: 0.0 Sum: 0.0	0-11.5 Mean: 5.0 Sum: 23.3	0-0 Mean: 0.0 Sum: 0.0	0-0 Mean: 0.0 Sum: 0.0	Range: 0:01-0:01 Mean: 0:01 Sum: 0:01	Range: 0-0 Mean: 0.0 Sum: 0.0
Effective praise	0.003-0:10 Mean: 0:02 Sum: 0:51	0.003-0:13 Mean: 0:01 Sum: 0:58	0:003-0:15 Mean: 0:03 Sum: 2:39*	0:004-0:08 Mean: 0:02 Sum: 0:20	Range: 0.3s-0:11 Mean: 0:03 Sum: 0:32	Range: 0.3-3:00 Mean: 0.8s Sum: 0:08
Neutral feedback	0.003-0:03 Mean: 0:02 Sum: 0:13	0.003-0:07 Mean: 0:03 Sum: 1:14	0:003-0:10 Mean: 0:02 Sum: 1:11	0:006-0:14 Mean: 0:03 Sum: 4:01*	Range: 0:01-0:04 Mean: 0:02 Sum: 0:08	Range: 0:02-0:22 Mean: 0:15 Sum: 1:16
Process feedback	0-0 Mean: 0 Sum: 0	0-0 Mean: 0 Sum: 0	0-0:05 Mean: 0:04 Sum: 0:08	0-0 Mean: 0.0 Sum: 0.0	Range: 0:05-0:17 Mean: 0:12 Sum: 0:50*	Range: 0-0 Mean: 0.0 Sum: 0.0
Echo	0.004-0:06 Mean: 0:01 Sum: 0:48	0.004-0:10 Mean: 0:02 Sum: 2:19	0:003-0:15 Mean: 0:01 Sum: 1:56	0:003-0:06 Mean: 0:01 Sum: 1:20	Range: 0.7s-0:03 Mean: 0:01 Sum: 0:12	Range: 0-0:02 Mean: 0:01 Sum: 0:02
Peer and self-correction	0:004-0:41 Mean: 0:04 Sum: 1:37*	0-7:30 Mean: 7:30 Sum: 7:30*	0:04-2:45 Mean: 0:34 Sum: 11:20*	0-0 Mean: 0.0 Sum: 0.0	Range: 0:06-1:15 Mean: 0:39 Sum: 5:52*	Range: 0-0 Mean: 0.0 Sum: 0.0
Personalization	0:004-3:22 Mean: 0:16 Sum: 22:16	0-0:16 Mean: 0:09 Sum: 2:52	0-0:13 Mean: 0:07 Sum: 0:43	0:03-0:20 Mean: 0:12 Sum: 0:38	Range: 1:04-1:04 Mean: 1:04 Sum: 1:04*	Range: 0-0 Mean: 0.0 Sum: 0.0
Group work	2:36-13:10 Mean: 7:00 Sum: 21:04*	1:30-18.30 Mean: 9:50 Sum: 53:40*	0-11:34 Mean: 1:18 Sum: 5:12*	0-0 Mean: 0.0 Sum: 0.0	Range: 0:28-0:48 Mean: 0:42 Sum: 3:21*	Range: 0-0 Mean: 0.0 Sum: 0.0
Pair work	0-0 Mean: 0.0 Sum: 0.0	0-0 Mean: 0.0 Sum: 0.0	0:06-1:55 Mean: 1:04 Sum: 6:28	0-25:40 Mean: 25:40 Sum: 25:40*	Range: 1:02-1:33 Mean: 1:10 Sum: 4:41*	Range: 0-0 Mean: 0.0 Sum: 0.0

Appendix 9
Lower achieving student interview transcripts

Question	AfL Students	Non-AfL Students
1. ¿Que te parecen las clases de ciudadanía?	**STU 1:** Siempre, cada semana, pues... no sé, nos lee un tema nuevo, y nos lee cuentos... de lo que... Y, y empezamos los temas, pues, con un cuento. **STU 2:** Mm... bien. Eh, pues, aprendo cosas. Y hacemos juegos a veces. Y también nos leen un cuento…. **STU 3:** Mmm... chulas, hombre, casi siempre trabajamos en grupo y me mola.	**STU 1:** Me gustan. Pues, nos digan %x…x% como aprender a ser mejor ciudadanos **STU 2:** No estan…no estan muy mal pero tampoco sean lo mejor **STU 3:** Están bien. explicamos muchas cosas. Y es muy divertido -
2. ¿Piensas que ciudadanía es fácil o difícil?	**STU 1:** Fácil . **STU 2:** Fácil. **STU 3:** ((Mimes)) No sé, normal. Pienso que... ni muy fácil ni muy difícil.	**STU 1:** Fácil **STU 2:** Ni fácil ni difícil. Es medio. **STU 3:** Yo pienso así-así, porque hay unos días que sea muy allá
3. ¿A veces cuando la clase de ciudadanía termine, te gustaría que dura un poquito mas?	**STU 1:** Algunas veces sí y otras… no. Porque hay veces que… que es un poco aburrida **STU 2:** Mm… a veces. Mm… cuando estamos jugando a juegos **STU 3:** Si. Porque a veces hay unas cosas que son muy entretenidas y que, como se va a acabar la hora dentro de poco, quiero que continúen.	**STU 1:** Un poco **STU 2**: Si **STU 3:** ((nods))
4. ¿te esfuerzas mucho en ciudadanía para que tu profesor es contento?	**STU 1:** Sí. **Stu 2:** Sí. **STU 3:** Sí.	**STU 1:** Bastante **STU 2:** Bueno, unas veces si, otras no. No se, porque a veces me aburro un poco **STU 3:** Sí.
5. ¿Crees que los temas de ciudadanía serian útil en el futuro? Para ti?	**STU 1:** Para mí sí. Porque en Ciudadanía nos explica cosas... Y si por ejemplo yo de mayor conduzco y me voy a un sitio donde hablan inglés, pues... les ent... **STU 2:** Algunos. Eh, como... eh... las actividades... que hace. Ah, mm las emociones **STU 3:** No sé, creo que sí.	**STU 1:** Sí. **STU 2:** Sí., Porque así sabemos comportar en la ciudad. **STU 3:** Yes. Porque, ahora mismo estamos dando democracia y todo esto, pues nos sirve para cuando seamos mayores.

6 ¿A ti te gusta aprender en Ingles?	**STU 1:** Sí. Porque... Porque, eh, las... porque las palabras son distintas y... : Y aprendo nuevas cosas. **STU 2:** Sí. Para... porque así puedo aprender más. **STU 3**: Sí Porque me gusta aprender en nuevos idiomas. No sé, porque para cuando sea mayor y me hablen, y tenga un amigo inglés, para no estar hablando en español.)) Si voy a Inglaterra o a London.	**STU 1:** Si, Porque alguna vez cuando vas a viajar a algún sitio, si hablas en Ingles, te entiendan mejor. **STU 2:** Depende de las clases. Las ingles, bueno, estan bien, pero plastica me gusta mas, también damos en ingles, también science y eso. **STU 3**: Si. Bueno me gusta mucho ingles pero me cuesta un montón aprenderlo, pero me gusta
7. prefieres aprender en español o en ingles?	**STU 1:** Español **STU 2:** En inglés. : Porque eh, hay otros lugares que se habla más en inglés. : Y para... para ir a lugares para que me entiendan. **STU 3:** En ingles. Pues porque me gusta saber más cosas en inglés. Porque en español ya me lo pueden enseñar mis padres.	**STU 1:** Yo prefiero en español, pero en ingles, hay cosas mejores **STU 2:** Español, Me gusta mas y entiendo mas las cosas **STU 3:** En Ingles. Porque así cuando sea mayor. Si tengo que viajar a un país y no sean, por ejemplo frances, si hablas frances, pero vamos… Algo que no, pues %x…x% ingles, y así todos me entienden si pido otro mejor
8. ¿crees que aprender en ingles es fácil o difícil?	**STU 1:** Sí. Pues es difícil porque como no sé mucho inglés todavía, pues... se me da un poco mal. **STU 2**: (dificil) Ehm... al principio. Pues cuando no sabes nada, pues te, te parece difícil. Pero cuando sabes ya, pues, es más fácil **STU 3**: Mmm... ((mimes)) Más que en español. ((Smiling)) Porque como yo nací en España, sé más español que inglés.	**STU1:** Normal. porque algunas palabras %x…x% difícil. **STU 2:** Ah, depende si no entiendes para %x…x% tampoco es muy difícil porque te las explica la profe. **STU 3:** No.
9. ¿piensas que ingles te va a servir en el futuro?	**STU 1:** Si, Porque si me voy a un sitio, pues... y pregunto a alguien dónde está no sé qué, pues me dicen... pues me lo, me lo, si hay un inglés, pues me lo dice en inglés y yo le entiendo. **STU 2:** Eh, sí. Porque así, ah... puedo... hablar inglés y... puedo hacer cosas útiles. **STU 3**: Si. Como para hacer nuevos amigos o para irme a Inglaterra o	**STU 1:** Pues sí. **STU 2:** : Sí. para cuando vaya a otros paises si me entienden **STU 3:** Porque quiero trabajar de científica, científica lo hacen en España, también en otro país

	algún restaurante inglés ((mimes))...	
10. ¿Crees que va bien tus clases este trimestre?	**STU 1:** Sí. **STU 2**: Sí. **STU 3:** Normal.	**STU 1:** Sí. **STU 2**: Sí. **STU 3:** Mmm, sí.
11. ¿piensas que vas a sacar buenas notas en las clases…en las clases que das en Ingles?	**STU 1**: Sí **STU 2:** Sí. **STU 3:** Mm, creo que si... más o menos Porque hay veces que se me olvida hacer los deberes y eso Rebeca lo cuenta.	**STU 1:** Sí. **STU 2:** Sí. **STU 3**: Bueno….((shakes hand indicating so-so)) pienso que así así, ¿eh? Porque en science, ya que siempre copiamos, me cuesta mucho estudiar cuando en mi familia, nadie sabe ingles.
12. ¿piensas que te van bien las clases de Ingles?	**STU 1:** Algunas cosas. Pues es que los adverbios no se me dan muy bien y... y entonces... pues eh, hicimos un examen y tuvimos que hacer... eh, una cosa de poner tres ejemplos de adjetivos, de... de nombres, de verbos, de adverbios... y, y como no sabía mucho de adverbios, pues no sabía qué poner. **STU 2:** Eh, sí. **STU 3:** Mm, sí.	**STU 1:** Sí. **STU 2**: Bueno, a veces si, a veces no. A veces no, porque a veces me %x…x% la profe y también por que…me aburro **STU 3:** Sí.
13. ¿crees que un dia serás capaz de hablar muy muy muy bien en Ingles?	**STU 1:** No. Porque es muy difícil. **STU 2:** Sí. **STU 3:** Mm, no sé. A lo mejor, si estudio mucho inglés, sí.	**STU 1:** Sí. **STU 2:** Sí. **STU 3:** Sí.
14. En las clases que das en Ingles, ¿normalmente entiendes lo que tienes que hacer, y como tienes que hacerlo?	**STU 1:** Sí. **STU 2:** Eh... bueno, a veces no. Porque, eh... dice algunas cosas y yo no las entiendo. Pero luego se las digo otra vez y me lo dice, otra vez **STU 3:** Sí, siempre. Aunque hable muy rápido, siempre.	**STU 1:** Sí. **STU 2:** Sí., a veces sí., a veces no. A veces no, porque no entiendo lo que %x…x% las instrucciones? **STU 3:** Sí…..son cosas como %x…x% me cuesta a veces entenderlo, pero casi…normalmente si.

15. Este trimestre, ¿tienes la impresión que estas aprendiendo mucho ingles?	**STU 1**: Sí. **STU 2**: Sí. **STU 3:** Sí.	**STU 1**: Sí. **STU 2:** Sí. **STU 3:** Yes
16., ¿Crees que vas a aprobar, aprobar las materias que imparten in Ingles?	**STU 1:** Pero no muy buena nota. No creo. Porque es que... eh... ((smiles and looks to the camera)) no sé... porque como son difíciles... **STU 2**: (no response) **STU 3**: Sí, creo que sí. No sé, porque hay algunas clases que tampoco es que me porte muy bien	**STU 1:** Sí. **STU 2:** Sí. **STU 3:** Yes
17. ¿Levantas la mano para hablar cuando estas en clase de Ingles?	**STU 1:** A veces. Porque hay algunas cosas que no las sé. **STU 2:** A veces. Porque a veces no sé las respuestas y otras, sí. **STU 3:** Sí.	**STU 1:** A Veces. Porque algunas veces, no me se bien la respuesta **STU 2**: A veces si, a veces no. En ciencias, casi siempre no levanto. En ingles, bueno, cuando…cuando…cuando me se la respuesta la levanto. En ciudadanía igual **STU 3:** A veces. : Porque hay preguntas que no se %x…x% desde el principio. Y las tengo que volver a repetir y estudiar mucho (en ciudadania) Porque hay algunas cosas que estoy escribiendo, haciendo apuntes y se me olvida levantar la mano y todo
18. si estudias mucho por una examen ¿crees que vas a mejorar tu nivel de ingles?	**STU 1:** Sí. **STU 2:** Sí. **STU 3:** Creo que sí.	**STU 1:** Pues, estudiando se aprenden mas, pero con otras cosas también **STU 2:** Sí. Porque así, se mejor lo que estabamos dando. Eh, eh, ortografía **STU 3:** Sí. Porque…a mi me gusta estudiar, es %x…x% mi, pues, porque mis padres me encierren en la habitación y yo me lo tengo que leer y ellos me preguntan.
19. ¿Te preocupes mucho si te equivocas cuando estas hablando en Ingles?	**STU 1:** Sí, porque algunas cosas, a lo mejor, eh... pues, al decir en vez de 'team', de un equipo, pues digo /'taim/, de hora. Y... y que, como la... y a veces me equivoco como son t... casi iguales, eh, 'mouse' y 'house', en vez de decir 'mouse' me equivoco y digo 'house' ((smiling)). **STU 2**: Eh... Sí... Porque, no sé, por si,	**STU 1:** Alguna vez…porque…no….es…no se, me preocupo no se **STU 2:** No **STU 3:** Sí. Ay, porque no estoy muy segura, es que me preocupa no, no se bien el ingles

	eh... pues ((mimes)) no sé ((smiles)). Porque... eh... A veces, eh... ((smiles)) es que... **STU 3:** Mm... no, no me preocupo mucho...	
20. ¿Tienes miedo que tus compañeros se ríen de ti cuando estas hablando en Ingles durante la clase?	**STU 1:** No. **STU 2**: No. **STU 3**: No.	**STU 1**: Pues, no **STU 2:** A veces. Porque me daria vergüenza si que se ríen todo de mi **STU 3:** No
21. ¿Te sientes mas nervioso en clases que das en Ingles que en las clases que das en español?	**STU 1:** Algunas veces sí y otras no. Porque algunas son tan... un poco difíciles y, y... y en Matemáticas, por ejemplo, le entiendo mejor. **STU 2:** Sí.. Porque en inglés estoy aprendiendo más, y en español, ya sé hablar en español. ((de no levanter su mano) Porque... porque... por si me equivoco o algo así. **STU 3:** Sí. Porque... Más nervioso porque cuando Rebeca me dice que salga a decir una historia o algo, ahí me atraganto mucho o... no sé qué pasa. Y en Lengua, por ejemplo, si tengo que hacer una historia en español...pues no me preocupo porque sé leer bien	**STU 1:** Un poquito mas, Porque es mas difícil aprender en Ingles **STU 2**: Bueno, no **STU 3**: No
22. ¿Te pones muy nervioso cuando tienes que hacer un examen en Ingles que un examen en español?	**STU 1:** Sí. Porque como no... son un poco... ((the camera falls down)) así que... que hay algunas cosas que no las entiendo y que... no estoy muy seguro de... de las cosas... **STU 2:** Sí. Porque algunas palabras no las entiendo y, entonces, pues me puedo equivocar. **STU 3:** Eh, ss... mm, igual. Porque si en inglés entiendo, eh, sé leer, y entiendo lo que dicen y en español. Es una pregunta igual, solo que la voy a entender mejor, pero van a estar igual, las voy a entender me... igual, las dos.	**STU 1:** Pues, un poco mas nervioso, Porque las preguntas en ingles y te cuestan mas saber que tienes que poner. **STU 2:** Sí. Porque a lo mejor estoy tan nervioso que fallo en algo. porque estoy tan nervioso que…que a lo mejor escribo rapido así, y fallo mas **STU 3:** Mmm…si es de ingles, si. Porque como se mira mal los verbos, me mira mal escribir, pues, me cuesta. Entonces, me pongo…ay, ¿que escribo? Entonces, me cuesta

23. ¿Prefieres hacer exámenes o actividades en clase?	**STU 1:** Actividades. Por ejemplo, hacer cosas en grupo o las parejas, que... o como ayer que nos pusieron unas palabras del spelling test y teníamos que hacer frases con ellas. **STU 2:** ((Smiling)) Actividades. Pues juegos, en grupo o... **STU 3:** Otras actividades. Eh, por ejemplo, como trabajar en grupos, hacer historias en clase, para que ella, cuando salgamos, nos corrija mejor.	**STU 1:** Actividades, Ejercicios. preguntas **STU 2:** Prefiero actividades. : Ejercicios. A veces hacemos juegos **STU 3:** Pues, actividades de ciudadanía, prefiero hacer actividades que están %x…x%, trabajar en grupos
24. ¿Estudiando para un examen te va a ayudar con tu Ingles?	**STU 1:** … **STU 2:** Sí. **STU 3:** Sí.	**STU 1:** Sí. **STU 2:** Sí. **STU 3**: Sí.
25. ¿Que actividad en particular te ayudaría mas con tu ingles?	**STU 1:** En parejas **STU 2:** Escribir. Porque la... hay que, que... hay letras que me confundo a veces para escribir. **STU 3:** Pues... escuchar y... escribir. Porque cuando nos manda escribir una historia, escribo, se la entrego, la corrige y ya veo mis fallos y los... mejoro.	**STU 1:** Hablar en grupos, con mis companeros. Em, repasar palabras que me cuestan **STU 2:** Cuando el profesor nos explica algo y en grupos **STU 3:** Hablar